

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

IDENTIFICATION DE NOUVEAUX GÈNES *CBF* CHEZ LE BLÉ HEXAPLOÏDE  
NORSTAR ET IDENTIFICATION DE POLYMORPHISMES AU NIVEAU DE  
GÈNES *CBF* CHEZ DES CULTIVARS DE BLÉ POSSÉDANT DES CAPACITÉS  
DIFFÉRENTES D'ACCLIMATATION AU FROID

MÉMOIRE DE RECHERCHE  
PRÉSENTÉ  
COMME EXIGENCE  
DE LA MAÎTRISE EN BIOLOGIE

PAR  
SARA MOHSENI-MASOULEH

JUILLET 2012

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Mes premiers remerciements s'adressent à mon directeur de recherche Dr. Jean Danyluk pour me guider et pour transmettre sa connaissance et les attentives relectures de chaque partie de mon mémoire.

À tous les professeurs du groupe: Dr. Fathey Sarhan, Dr. Normand Chevrier, Dr. Mario Houde, Dr. François Ouellet et la chercheuse Dre. Zahra Agharbaoui.

À mes collègues du laboratoire Zakia Djillali, Joseph Nankeu, Khalil Kane, Amadou Diallo, Mohamed Ali Ali Benali, Mélanie Grondin et Hua Che pour leur aide indispensable.

Je remercie tous les membres de ma famille, pour leur soutien et leurs encouragements.

## TABLE DES MATIÈRES

REMERCIEMENTS.....	ii
LISTE DES FIGURES.....	v
LISTE DES TABLEAUX.....	vi
LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES .....	vii
RÉSUMÉ .....	ix
INTRODUCTION .....	1
CHAPITRE I	
REVUE DE LITTÉRATURE.....	4
1.1 Stress au froid et développement végétal .....	4
1.2 Dommages causés par le gel.....	6
1.3 Mécanisme d'acclimatation au froid .....	7
1.4 Gènes COR.....	8
1.4.1 AFP (Anti Freeze Protein) .....	9
1.4.2 Protéines LEA (Late Embryogenesis Abundant).....	10
1.4.3 Autres gènes COR.....	10
1.5 Les rôles des autres facteurs de régulation de l'acclimatation au froid.....	11
1.5.1 Le rôle du calcium.....	11
1.5.2 Le rôle de l'ABA (acide abscissique) .....	12
1.6 Facteur de transcription AP2 .....	12
1.7 Gènes CBF chez Arabidopsis .....	13
1.7.1 Gènes CBF chez l'orge .....	15
1.7.2 Gènes CBF chez le blé .....	18
1.8 Vernalisation et sa relation avec la résistance au gel.....	19
1.9 Rôles possibles d'une amplification du contenu génique.....	20
1.10 QTL (Quantitative trait locus) .....	21



1.10.1	Identification de QTL chez les céréales sur des gènes associés à d'autres stress que le froid .....	23
1.10.1.1	QTL chez les céréales .....	23
1.10.1.2	Identification des QTLs de gènes CBFs chez le blé et l'orge .....	24
CHAPITRE II		
MATÉRIELS ET MÉTHODES.....		30
2.1	Identification de nouveaux gènes CBF du blé.....	30
2.2	Isolation des gènes CBF par PCR.....	30
2.3	Analyses phylogénétiques et bioinformatiques .....	32
2.4	Identification de polymorphismes au niveau des gènes CBF chez deux cultivars de blé printemps .....	35
2.4.1	Génération d'amorces spécifiques et PCR.....	35
2.4.2	Clonage et analyses bioinformatiques.....	36
CHAPITRE III		
RÉSULTATS .....		38
3.1	Identification de nouveaux gènes CBF.....	38
3.2	Analyse bioinformatique des protéines CBF .....	42
3.3	Identification de polymorphismes chez les cultivars de printemps Chinese Spring et Manitou .....	54
CHAPITRE IV		
DISCUSSION .....		72
ANNEXE		
LISTE DES FIGURES ET TABLEAUX SUPPLÉMENTAIRES .....		80
RÉFÉRENCES.....		142

## LISTE DES FIGURES

Figure	Page
3.1 La relation phylogénétique entre les gènes CBF des Triticeae. ....	41
3.2 La relation phylogénétique entre les gènes CBF des groupes IIIa/b/c/d chez les Triticeae. ....	43
3.3 La relation phylogénétique entre les gènes CBF des groupes IVa/b/c/d chez les Triticeae. ....	44
3.4 L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine AP2 et les signatures avoisinante des protéines CBF des Groupes I et II d'orge ( <i>Hordeum vulgare</i> ), blé ( <i>Triticum spp.</i> ) et seigle ( <i>Secale cereale</i> ). ....	47
3.5 L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine C-terminal des protéines CBF des Groupes I et II d'orge ( <i>Hordeum vulgare</i> ), blé ( <i>Triticum spp.</i> ) et seigle ( <i>Secale cereale</i> ). ....	48
3.6 L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine AP2 et les signatures avoisinante des protéines CBF des Groupes III d'orge ( <i>Hordeum vulgare</i> ), blé ( <i>Triticum spp.</i> ) et seigle ( <i>Secale cereale</i> ). ....	49
3.7 L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine C-terminal des protéines CBF des Groupes III d'orge ( <i>Hordeum vulgare</i> ), blé ( <i>Triticum spp.</i> ) et seigle ( <i>Secale cereale</i> ). ....	50
3.8A L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine AP2 et les signatures avoisinante des protéines CBF des Groupes IV d'orge ( <i>Hordeum vulgare</i> ), blé ( <i>Triticum spp.</i> ) et seigle ( <i>Secale cereale</i> ). ....	51
3.8B L'analyse comparative des noyaux hydrophobiques (HC) présent entre le domaine AP2 et la région C-terminale des protéines CBFIVd-9 et IVd-22. Cette figure est la suite de la figure 8 A. ....	52
3.9 L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine C-terminal des protéines CBF des Groupes IV d'orge ( <i>Hordeum vulgare</i> ), blé ( <i>Triticum spp.</i> ) et seigle ( <i>Secale cereale</i> ). ....	52

## LISTE DES TABLEAUX

Tableau	Page
2.1 La nomenclature et les caractéristiques de nouveaux gènes chez le blé hexaploïde. ....	34
3.2 Liste des homéologues CBF qui contiennent des mutations qui affectent la séquence/structure consensus CBF. ....	55
3.3 Liste de tous les gènes/paralogues/haplotypes de CBF retrouvés chez le blé hexaploïde. ....	57
3.4 Nombre de polymorphismes (SNP) causant des mutations faux- et non-sens sur les différentes parties de la protéine CBF. ....	67
3.5 Liste des mutations survenant dans des portions conservées de gènes CBF groupées selon la région impliquée. ....	70

## LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

ABA	Acide abscissique
ABRE	Absciscic acid responsive element
AFP	Anti Freeze Protein
ALMT1	Le gène de résistance au stress d'Al
AP2	APETALA2
CAX1	Calcium Exchanger 1
CDP	C-repeat Binding Factor
COR	Cold-regulated
CRT	C-repeat
DRE	Drought responsive element
DREBI	DRE Binding proteins I
ERD	Early dehydration-inducible
EREBP	Ethylene Responsive Element Binding Protein
HAT	Histone acetyltransferase
HC	Hydrophobic clusters
HCA	Analyse des clusters hydrophobiques
LCQ	Locus de caractères quantitatifs
LEA	Late Embryogenesis Abundant
KIN	Cold-induced

LTI	Low temperature-induced
PLC	Phospholipase C
PLD	Phospholipase D
QTL	Quantitative trait locus
RD	Responsive to desiccation
SNP	Single-Nucleotide Polymorphisms

---

## RÉSUMÉ

Les *CBFs* représentent une famille de gènes impliqués dans le mécanisme de résistance de la plante au gel. Chez le blé hexaploïde, Badawi et ses collègues (2007) ont trouvé 37 gènes *CBFs* qui représentent 15 groupes de gènes différents avec 1 à 3 copies. Ces gènes ont été classifiés en 10 sous-groupes différents en fonction de la structure secondaire de leurs protéines. Les analyses phylogénétiques et bioinformatiques révèlent que le blé hexaploïde pourrait avoir plus de 25 groupes de gènes *CBF*. Nous avons poursuivi l'isolation de gènes *CBF* afin d'identifier le répertoire complet de cette famille chez le blé hexaploïde tolérant au gel, Norstar. Nous avons identifié 65 gènes *CBF* chez le blé. L'analyse phylogénétique des séquences nucléotidiques codant pour le domaine AP2 et la signature *CBF* révèlent que le blé contiendrait 27 groupes de gènes (avec 1 à 3 homéologues chacun provenant des 3 génomes du blé hexaploïde). L'analyse de leur structure secondaire montre qu'il est possible de classer les protéines *CBF* en plus de 10 sous-groupes pouvant posséder des fonctions différentes ou complémentaires. De plus, la séquence/structure de plusieurs copies (homéologues) *CBF* chez le blé hexaploïde dévie de celle conservée dans les autres membres du groupe suggérant que ces protéines peuvent représenter des produits de gènes sub optimaux. Ceci pourrait avoir un impact sur l'activité de ces *CBF* et affecter la régulation des gènes *COR* et le développement de la tolérance au gel. En tout, Norstar possède 2 pseudogènes *CBF* exprimés et 24 des 64 protéines analysées peuvent être considérés comme sub optimales.

La présence d'un QTL a déjà été associée avec des gènes *CBF* chez le blé et l'orge. Comme les *CBF* peuvent être à la base de cet important phénotype, une hypothèse plausible pour expliquer leur contribution pourrait être que les gènes *CBF* possèdent des différences moléculaires au niveau des régions codantes qui influencent l'activité biologique des *CBF* exprimés. Afin de vérifier cette possibilité, nous avons comparé les régions codantes des gènes *CBF* du cultivar d'hiver Norstar à celles de deux cultivars de printemps, Chinese Spring et Manitou. Les résultats montrent que les deux cultivars de printemps possèdent aussi deux pseudogènes *CBFIIIc-13.1* et *CBFIVd-B22* mais le contenu et la nature des gènes optimaux et sub optimaux varient. Parfois le gène *CBF* chez le cultivar de printemps possède une séquence qui est plus proche d'une séquence consensus que celle provenant du cultivar d'hiver Norstar comme observé pour *CBFIIIc-24.2b*, et parfois la séquence dévie du consensus comparativement à celle du cultivar Norstar comme observé pour *TaCBFIVb-20b*. La présence d'haplotypes différents chez le blé suggère que certains allèles peuvent coder pour des protéines *CBF* plus optimales au niveau de leur activité biologique et d'autres moins optimales. Il sera important d'étudier l'activité

des protéines codées par ces haplotypes afin d'évaluer les variantes qui seraient les plus aptes à être combinés dans un génotype afin d'augmenter la régulation positive du développement de la tolérance au gel. Cette information pourra être utilisée par les sélectionneurs lors de croisements de divers génotypes de blé afin de produire des plantes plus tolérantes au gel.



## INTRODUCTION

La croissance de la population a augmenté de manière significative depuis la révolution industrielle du début du 18<sup>ième</sup> siècle. Une publication du bureau de recensement des États-Unis (United State Census Bureau) indique que la population du monde a dépassé 6.5 milliards d'individus le 25 février 2006. En 2007, l'United Nations Population Division a prédit que la population du monde atteindra 10 milliards en 2055. Avec une augmentation de population d'à peu près 90 millions par année, la demande pour l'alimentation et l'énergie devrait s'intensifier. Cependant dans les régions nordiques, le froid est un facteur limitant pour la production agricole. En considérant que les plantes sont des organismes sensibles et ne peuvent pas contrôler leur propre température, la tolérance au froid est très importante dans les climats tempérés pour la survie des cultures. Le Canada est un des plus grands exportateurs de blé au monde. C'est une ressource importante sur le plan économique et sur le plan nutritif. Le blé et plusieurs autres espèces des climats tempérés ont développé une réponse adaptative, l'acclimatation au froid, qui leur permet de supporter le gel (Catala et Salinas, 2008).

Plusieurs gènes sont induits durant l'acclimatation au froid et sont impliqués dans le développement de la tolérance au gel. Ces gènes peuvent être classés en deux différents groupes. Le premier groupe fonctionne directement pour protéger la plante contre un stress environnemental (LEA, AFPs, COR) et l'expression des différents membres de ce groupe a été mainte fois corrélée avec la capacité des plantes à tolérer le gel. Le deuxième groupe comprend des régulateurs de l'expression des gènes et de la transduction du signal du stress. Dans ce deuxième groupe, une des familles de gènes très importante pour le processus d'acclimatation au froid est la famille *CBF*. Son étude chez les céréales sensibles et tolérantes au froid représente un excellent modèle d'étude comparative pour comprendre la fonction et la régulation de ces gènes de tolérance aux basses températures. Chez les *Poaceae*, plusieurs gènes *CBFs* ont été isolés et caractérisés chez les espèces résistantes au gel, telles que l'orge (*Hordeum vulgare*), le blé (*Triticum spp.*), le seigle (*Secale cereale*), le ray-grass (*Lolium tremulentum*), ainsi que chez des

cultures sensibles comme le riz (*Oryza sativa*) et le maïs (*Zea mays*). En 2007, Badawi et ses collègues ont identifié 37 gènes *CBFs* représentant 15 groupes de gènes (avec 1 à 3 copies) chez le blé hexaploïde et ils les ont classifiés dans 10 sous-groupes I, II, IIIa, IIIb, IIIc, IIId, IVa, IVb, IVc et IVd en analysant la structure secondaire de leurs protéines. Les études sur l'expression et la phylogénie des différents gènes *CBF* montrent que les groupes IIIc, IIId, IVa, IVb, IVc, IVd sont apparus tardivement au cours de l'évolution des *Poaceae* et que leur expression est associée à la résistance au gel. L'étude de l'expression des gènes par RT-PCR a montré que cinq groupes de *CBFs* (IIId, IVa, IVb, IVc and IVd) sont plus fortement exprimés à de basses températures chez le blé tolérant au gel que chez le blé sensible (Badawi et al. 2007). Ces résultats suggèrent que ces groupes joueraient un rôle prépondérant dans le développement de la tolérance au gel. De plus, des analyses phylogénétiques et bioinformatiques des gènes *CBFs* présents chez les différentes espèces de blé ainsi que chez d'autres monocotylédones montrent qu'il existe probablement plus de 25 groupes de gènes *CBF* chez le blé hexaploïde. Puisque l'information sur la famille *CBF* chez le blé était incomplète, notre premier objectif était de poursuivre l'isolation de gènes *CBF* afin d'identifier le répertoire complet chez le blé hexaploïde tolérant au gel Norstar. Cette information est importante pour initier des études fonctionnelles pour comprendre leur rôle dans le développement de la tolérance au gel.

Les gènes *CBF* ont été associés avec un QTL de tolérance au gel chez le blé diploïde (Miller et al. 2006; Knox et al. 2010), le blé hexaploïde (Båga et al. 2007) et l'orge (Skinner et al. 2006; Francia et al. 2007). L'étude sur le QTL de *T. monococcum* indique que les polymorphismes détectés dans les gènes *TmCBF* de deux génotypes sont l'explication la plus probable pour les différences dans leur tolérance au gel. Chez l'orge d'hiver et de printemps, ce sont les différences dans le nombre de copies de gènes *CBF* fonctionnels, le nombre de pseudogènes, et aussi le changement polymorphique qui peuvent être les causes de différences dans la tolérance au gel. Donc les deux hypothèses plausibles pour expliquer la contribution des *CBF* au QTL pourraient être qu'ils possèdent : 1) des différences moléculaires au niveau des régions régulatrices d'expression (promoteurs) ou au niveau génomique (duplications de régions et augmentation du nombre de copies) qui induisent des expressions plus élevées de certains

*CBF* et/ou; 2) des différences moléculaires au niveau des régions codantes qui influencent l'activité biologique des *CBF* exprimés. Nous savons que l'expression des gènes *CBF* de plusieurs groupes est associée au développement de la tolérance au gel (Badawi et al. 2007), et l'isolation et l'analyse des promoteurs *CBF* présent chez le blé a été initié (Badawi et al. 2008). Toutefois, pour répondre clairement à cette première possibilité, il est nécessaire de cloner tous les promoteurs correspondant aux gènes identifiés chez le blé hexaploïde. Ce type d'expérimentation est laborieux et donc trop ambitieux comme suite à notre projet. Comme l'information entourant la région codante est devenue disponible suite à la première partie de notre travail, nous avons initié l'examen des régions codantes de gènes *CBF* chez le cultivar d'hiver Norstar et chez deux cultivars de printemps, Chinese Spring et Manitou afin de mettre en évidence les différences qui pourraient être à la base de la tolérance au gel plus élevée chez le blé d'hiver. Cette étude permettra de dresser une liste d'haplotypes *CBF* présents chez le blé et d'initier des études examinant l'activité biologique de ces variantes. L'information fournie par ces études pourra être utilisée par les sélectionneurs lors de croisements de divers génotypes de blé afin de produire des plantes plus tolérantes au gel.

## **CHAPITRE I**

### **REVUE DE LITTÉRATURE**

La production des récoltes n'est pas seulement pour l'alimentation, il y a quelques années que les scientifiques ont découvert la production des biocarburants. Ces derniers sont produits à partir des ressources renouvelables (plantes, ordures organiques, etc.) et peuvent être utilisés comme une alternative aux combustibles fossiles (gaz, pétrole). L'éthanol et le biodiesel sont les deux principaux biocarburants utilisés aujourd'hui. À cause du lien entre l'agriculture et la production d'énergie, l'industrie agricole joue un rôle majeur dans l'adoption et l'expansion des technologies d'énergie renouvelable. Alors, l'augmentation de la population implique un besoin alimentaire et énergétique accru, d'où la nécessité d'accroître la production (Thomashow, 2001). Toutefois, à cause de stress abiotiques, nous perdons une partie de nos récoltes chaque année.

L'étude de certaines plantes ayant la capacité de développer la tolérance aux stress tels que le gel, la déshydratation et la salinité va aider à améliorer et à stabiliser la productivité agricole lors d'exposition à ces stress (Thomashow, 2001). Un des grands stress environnementaux est le gel qui peut endommager les plantes et réduire les rendements causant un grand problème social et économique.

#### **1.1 Stress au froid et développement végétal**

Les plantes affrontent différentes conditions de stress environnementaux et elles ont développé des mécanismes pour tolérer ces genres de conditions via plusieurs adaptations physiques. Les températures extrêmes, la sécheresse et le taux élevé de salinité sont les principaux stress qui affectent le plus la croissance des plantes et la productivité agricole (Xiong et al. 2002). Les plantes peuvent vivre seulement dans des conditions (ex. température) spécifiques pour leurs espèces. Leurs performances diminuent avec le froid, un stress abiotique courant que les plantes doivent affronter pendant leur vie (Bracale et Coraggio, 2003). Le stress

au froid comprend les températures froides ( $<20^{\circ}\text{C}$ ) et le gel ( $<0^{\circ}\text{C}$ ). Le froid limite la distribution géographique des plantes et diminue la productivité en réduisant la période de développement (Sharma et al. 2005).

Une distribution géographique restreinte de plantes a souvent obligé les gens à cultiver les plantes à l'extérieur de leurs limites naturelles. Plusieurs de ces espèces sont les plantes tropicales et subtropicales, et elles résistent difficilement au froid (Bracale et Coraggio, 2003). Il existe plusieurs plantes d'importance économique tel le soya, le coton, le maïs, le riz, la tomate, la banane et la mangue qui sont sensibles au froid (Sharma et al. 2005). Elles ne sont pas capables de s'adapter au froid et ne tolèrent pas le gel. La membrane plasmique a une structure fluide, mais le stress au froid diminue la fluidité de la membrane et perturbe son fonctionnement (Chinnusamy et al. 2006).

Le stress au froid réduit la croissance de plantes avec un effet négatif et imprévisible sur la biomasse qui cause un grand intervalle entre le potentiel et la productivité actuelle. Aussi, le stress au froid a un effet sur la qualité des produits. Il diminue la synthèse, l'accumulation et l'entreposage de protéines et polysaccharides. De plus, la production de fruits est affectée par la diminution de température, qui réduit la valeur nutritive et le goût du produit (Bracale et Coraggio, 2003).

La sensibilité au froid varie en fonction des espèces et en fonction des cultivars de la même espèce. Par contre, la tolérance ne dépend pas juste de la résistance de la plante, mais dépend aussi du taux de refroidissement et de la présence d'autres conditions environnementales, comme l'humidité, le vent (qui peut causer une sécheresse), ou l'intensité lumineuse (qui peut causer une photo inhibition) (Bracale et Coraggio, 2003).

La sensibilité des plantes aux basses températures peut être facilement classée. Les plantes qui sont endommagées par le froid (entre  $0$  et  $15^{\circ}\text{C}$ ), sans être exposées au gel, sont les plantes sensibles au froid. Ces plantes meurent ou subissent des dommages suite à leur exposition au

froid. Les plantes tolérantes au froid, capables de résister à la formation de glace dans leurs tissus sont les plantes tolérantes au gel. Ces plantes deviennent plus résistantes aux blessures causées par le stress au froid après une période d'acclimatation au froid. Parmi elles on compte la majorité des plantes tempérées tel le blé, l'orge, l'avoine et le seigle. Elles sont capables de résister au froid et de développer une tolérance à la formation de glace extracellulaire (Chinnusamy et al. 2006).

## 1.2 Dommages causés par le gel

Quand la température diminue à moins de 0°C, la glace se forme dans l'espace intercellulaire du tissu de la plante, car le point de congélation du liquide intercellulaire est plus haut que celui du liquide intracellulaire. Quand la glace se forme dans l'espace intercellulaire, le potentiel de l'eau à l'extérieur de la cellule diminue. Alors on trouve un mouvement de l'eau liquide, de l'intérieur de la cellule vers l'extérieur (Thomashow, 1998). Les cristaux de glace commencent à se former lorsque la température atteint -2°C. Pendant la congélation, l'eau sort de la cellule peu à peu. Le niveau de déshydratation augmente parallèlement avec la diminution de la température (Yamada et al. 2002).

Yamada et ses collègues (2002) ont examiné les réponses de cellules de plantes à une baisse de température dans les groupes de plantes qui sont sensibles (violette africaine) et résistantes au gel (orchidées). Les résultats montrent que chez les plantes sensibles au froid, la membrane plasmique ainsi que leur paroi n'est pas capable d'empêcher le cheminement de glace de l'extérieur vers l'intérieur de la cellule; par contre chez les plantes résistantes au gel, leur membrane et paroi possède cette propriété.

La formation de la glace à l'intérieur de la cellule est létale, car le grossissement des cristaux de glace intracellulaires perturbe l'intégrité du protoplasme vivant jusqu'à perforer la membrane plasmique. La diminution rapide de température peut causer la création de cristaux de



glace intracellulaire. Quand la température baisse graduellement, la formation de cristaux de glace se produit dans la région intercellulaire. Les molécules d'eau migrent vers la surface des cristaux de glace et gèlent. La glace extracellulaire impose un stress sévère et une tension aux cellules, tissus et à la plante entière. La sortie de l'eau vers la glace extracellulaire a une influence profonde sur le volume de la cellule, son osmolarité, la concentration d'ions et l'interaction entre la membrane et paroi cellulaire surtout aux points d'attache (Guy, 2003).

Malgré que les dommages de la membrane plasmique soient principalement causés par la déshydratation cellulaire, des facteurs additionnels peuvent contribuer au dommage membranaire. La glace intercellulaire peut se coller à la paroi cellulaire et la membrane, et causer une rupture. La dénaturation des protéines qui survient lors de la diminution de température chez les plantes peut aussi endommager la plante (Thomashow, 1999).

### 1.3 Mécanisme d'acclimatation au froid

L'acclimatation au froid survient lors de la croissance aux basses températures (moins de 10°C) et permet d'induire une tolérance au gel et ainsi de minimiser les dommages causés par des températures moins favorables (Sharma et al. 2005). Chez un grand groupe de plantes tempérées, la croissance aux basses températures pour une période de temps, leur permet d'augmenter la tolérance contre le gel (Thomashow, 1999).

L'acclimatation au froid est associée à plusieurs modifications biochimiques, physiologiques, et moléculaires. Un des effets de l'acclimatation au froid comprend la stabilisation des membranes contre le gel. Plusieurs mécanismes sont impliqués dans la stabilisation des membranes comme les changements dans la composition lipidique de la membrane telle l'augmentation de la désaturation des acides gras dans la membrane phospholipidique et le changement au niveau des types de stérols. De plus, l'accumulation de saccharose et d'autres molécules de sucres simples qui surviennent pendant l'acclimatation au



froid pourraient participer à la stabilisation de la membrane, car ces molécules peuvent protéger les membranes *in vitro* contre le dommage induit par le froid.

Les cellules de plantes perçoivent le stress au froid à travers des changements de la fluidité de la membrane (Chinnusamy et al. 2006). Le froid augmente l'ion  $\text{Ca}^{2+}$  au cytosol. L'amplification du signal de  $\text{Ca}^{2+}$  et phospholipide peut être impliqué dans la signalisation au froid. En réponse à une exposition au froid, il a été montré que la phospholipase D (PLD) et la phospholipase C (PLC) sont activées simultanément. Vergnolle et ses collègues (2005) ont identifié 58 gènes inhibés via l'activité PLC et 87 gènes régulés via rétrogradation de PLD. La diminution de la température affecte l'absorption de l'eau et des substances nutritives, la fluidité membranaire et la conformation des protéines et des acides nucléiques. Le froid influence radicalement le métabolisme de la cellule, soit directement en réduisant le taux de réaction biochimique, soit indirectement à travers une reprogrammation de l'expression des gènes (Chinnusamy et al. 2006).

Le mécanisme de tolérance au gel empêche la dénaturation de protéines causées par le gel et la précipitation des macromolécules et diminue les dommages physiques qui sont causés par l'accumulation de glace extracellulaire (Thomashow, 1998). La tolérance contre le gel dépend du degré de développement du tissu et du type d'organe. Par exemple les grains sont très résistants au gel, par contre la racine est très sensible. Au niveau de la reproduction, la maturation du pollen est un stade très sensible au froid et au gel (Bracale et Coraggio 2003).

#### 1.4 GÈNES *COR*

Guy et ses collègues (1985) ont montré qu'un changement dans l'expression des gènes se produit au cours de l'acclimatation au froid et ainsi débute les études pour identifier et caractériser les gènes modulés par le froid. L'hypothèse était que quelques gènes sont

probablement impliqués dans la tolérance au gel et que l'étude de leurs fonctions et de leur régulation pouvait fournir une meilleure compréhension du processus d'acclimatation.

Les gènes qui répondent au froid peuvent être classés en deux différents groupes. Le premier groupe fonctionne directement pour protéger la plante contre un stress environnemental (LEA, AFPs). Le deuxième groupe comprend des régulateurs de l'expression des gènes et de la transduction du signal du stress (Shinozaki et Yamaguchi-Shinozaki, 2000). Les gènes modulés par le froid ont le plus souvent été nommés gènes *COR* (cold-regulated), mais aussi *LTI* (low temperature induced), *KIN* (cold-induced), *RD* (responsive to desiccation) and *ERD* (early dehydration-inducible) (Fowler et Thomashow, 2002)

Les recherches de Baker et ses collègues (1994) et Yamaguchi-Shinozaki et Shinozaki (1994), sur les promoteurs de gènes *COR* activés par les basses températures et le stress hydrique chez *Arabidopsis* ont permis d'identifier l'élément de régulation, le motif CRT/DRE (C-repeat/drought responsive element), qui a une séquence consensus de CCGAC. Les activateurs de transcription qui se lient au CRT/DRE ont été nommés CBF/DREBI (C-repeat Binding Factors/DRE Binding proteins I). Ces protéines possèdent un domaine de liaison à l'ADN de type AP2 (APETALA2).

#### 1.4.1 AFP (ANTI FREEZE PROTEIN)

Les protéines AFPs inhibent la croissance des cristaux de glace et la recristallisation. Elles sont efficaces dans la protection contre les dommages causés par le gel. Les AFPs ont été identifiées dans plusieurs organismes vivants pendant l'hiver comme les poissons tolérants, les arthropodes terrestres, les plantes vasculaires et non vasculaires, les champignons et les bactéries. Cependant leur structure et leur composition varie. L'hystérésis thermique (différence entre le point de fusion et le point de congélation de l'eau) varie entre 0,2 et 1,6°C pour les AFPs des végétaux, entre 1 et 2°C pour les AFPs des poissons et entre 5 et 10°C pour les AFPs des

insectes (Badawi, 2008). La fonction des AFPs de Type I est bien caractérisée. La molécule d'AFP s'adsorbe sur le treillage de glace par des liaisons hydrogène via des résidus de thréonines alignées qui sont situés sur la surface de la molécule (Wallis et al. 1997). Ici, elle crée une barrière pour l'eau qui diminue le taux de croissance des cristaux de glace. Cette liaison glace-protéine a été confirmée par cristallographie au rayon X. L'AFP inhibe aussi la recristallisation de la glace et cet effet peut protéger le tissu contre le dommage à une concentration de protéines 1000X moins que la concentration nécessaire pour inhiber initialement la croissance des cristaux de glace (Wallis et al. 1997).

#### 1.4.2 PROTÉINES LEA (LATE EMBRYOGENESIS ABUNDANT)

Les protéines LEA sont des polypeptides synthétisés à la fin de l'embryogenèse, et elles ont un rôle dans la tolérance à la dessiccation. Les gènes *LEA* ont été identifiés il y a 20 ans chez le coton et le blé, mais leurs fonctions exactes ne sont pas encore connues. Les gènes *LEA* ont été clonés dans plusieurs espèces de plantes et basés sur les motifs de leurs protéines correspondantes, elles ont été classifiées en au moins six différents groupes. Plusieurs protéines induites par le froid, le stress osmotique et l'acide abscissique exogène possèdent des similitudes avec les protéines LEA. Ces protéines pourraient protéger les structures cellulaires contre les effets d'une perte d'eau en fonctionnant comme un tampon d'hydratation, par isolation d'ions, ou par protection directe des autres protéines ou membranes (Guy, 2003).

#### 1.4.3 AUTRES GÈNES *COR*

Le gène *COR15a* code pour un polypeptide de 15kDa qui s'accumule dans le chloroplaste sous forme *COR15am* (mature). C'est un polypeptide hydrophile, et reste soluble jusqu'à ébullition. Ce polypeptide est composé d'un motif de 13 acides aminés répétés 4 fois. *COR15am* améliore la tolérance au gel chez des plantes transgéniques. Plusieurs types de gènes *COR* sont exprimés en même temps que *COR15a*. L'expression de *COR15a* a un effet dans la cellule

entière pas uniquement dans le chloroplaste. Steponkus et ses collègues (1998) ont proposé que le polypeptide mature COR15am retarde la formation de la structure lipidique hexagonale II qui survient durant la congélation des membranes plasmiques et chloroplastiques tandis que d'autres gènes *COR* codent pour des protéines qui participent à la stabilisation de membranes.

## 1.5 LES RÔLES DES AUTRES FACTEURS DE RÉGULATION DE L'ACCLIMATATION AU FROID

### 1.5.1 LE RÔLE DU CALCIUM

Pendant l'acclimatation au froid, l'ion calcium se fait le médiateur de la transduction du signal et est impliqué dans la phosphorylation de protéines et l'activation de plusieurs facteurs de transcription. L'augmentation du calcium cytosolique a lieu dans les minutes après l'exposition au froid suggérant son implication dans l'acquisition de la tolérance au gel (Sung et al. 2003). La diminution du calcium cytosolique par la séquestration dans les vacuoles à cause de la surexpression de CAX1 (Calcium Exchanger 1) rend des plantes transgéniques plus sensibles au froid (Sung et al. 2003).

Le calcium transmet le signal de température à la kinase MAPK (Mitogen Activated Protein Kinase). La cascade de MAPK est une importante partie de la voie de transduction du signal chez les plantes qui fonctionne de façon omniprésente durant la réponse à plusieurs signaux externes. Les MAPK deviennent actives lorsque les plantes sont exposées à un traitement hormonal, à des pathogènes, au froid, à la sécheresse et au stress salin (Sung et al. 2003). De plus, la MAPK activée par le choc de chaleur (HAMK) est différente de celle activée par le froid (SAMK). L'activation de HAMK et SAMK est déclenchée par le biais d'un changement dans la fluidité de la membrane qui est couplé avec l'augmentation du calcium cytosolique et l'action de protéines kinases dépendantes du calcium (CDPK, Calcium Dependent Kinases) (Sung et al. 2003).

### 1.5.2 Le rôle de l'ABA (acide abscissique)

Chen et ses collègues (1983) ont montré que le taux d'ABA augmente transitoirement au froid chez *Solanum commersonii*, une plante qui résiste au gel. En plus, ils ont découvert qu'une application exogène d'ABA augmente la tolérance au gel chez *S. commersonii*. Ces résultats ont guidé Chen et ses collègues (1983) à émettre l'hypothèse que l'acclimatation au froid est activée via l'action de l'ABA. Ils ont proposé que le froid provoque une augmentation du contenu en ABA qui déclenche l'activation du mécanisme de tolérance au froid. Chen et ses collègues (1983) ont confirmé ces résultats chez plusieurs autres plantes.

Quelques gènes *COR* (*COR78*, *COR47*, *COR6.6*, et *COR15a*) ont une très forte expression en réponse à l'application exogène d'ABA. L'activation de ces gènes par l'ABA est corrélée avec l'augmentation de la tolérance au gel. Les études sur les gènes *RD29a/COR78/LTI78* ont montré que les promoteurs de ces gènes contiennent un ABRE (abscisic acid responsive élément) et un DRE/CRT (Drought Responsive Element/C-Repeat). Ces éléments cis fonctionnent respectivement dans la réponse à un stress par une voie dépendante de l'ABA ou indépendante (Thomashow, 1998).

## 1.6 FACTEUR DE TRANSCRIPTION AP2

Une grande portion du génome des plantes code pour des facteurs de transcription: plus de 1600 gènes (6% du génome) ont été identifiés chez *Arabidopsis thaliana* (Riechmann et Meyerowitz, 1998) et chez le riz (Goff et al. 2002). Chez les eucaryotes, les facteurs de transcription sont classés en plusieurs familles, par rapport au type de domaine de liaison à l'ADN. Les facteurs de transcriptions AP2/EREBP (APETALA2/Ethylene Responsive Element Binding Protein) jouent un rôle important dans le contrôle du développement floral et la réponse aux stress biotiques et abiotiques.



Le génome d'*Arabidopsis* code pour 145 protéines appartenant à la famille AP2 qui partagent un domaine de liaison à l'ADN d'environ 60 acides aminés. Basé sur la similitude de leur domaine de liaison à l'ADN, ces protéines sont classées dans cinq sous-familles : AP2 (participe au développement), RAV, DREB (réponse aux stress abiotiques), ERF (réponse aux stress biotiques), et les autres qui ne sont pas classés. Les sous-familles DREB/CBF et ERF comprennent 125 gènes qui ont un seul domaine de liaison à l'ADN. Parmi eux, 121 gènes contiennent un motif conservé (WLC) au milieu de leur domaine AP2 (Sakuma et al. 2002).

### 1.7 GÈNES CBF CHEZ ARABIDOPSIS

Le groupe CBF/DREB1 chez *Arabidopsis* est composé de six paralogues, mais seuls trois CBF [CBF1 (*DREB1b*), CBF2 (*DREB1c*) et CBF3 (*DREB1a*)] sont induits au froid (Sakuma et al. 2002). CBF1, 2 et 3 sont présents sur le chromosome 4 et sont rapidement induits (dans les 15 min) en réponse aux basses températures, atteignant une expression maximale après 2 h d'exposition au froid (Gilmour et al. 1998; Liu et al. 1998; Stockinger et al. 1997; Zarka et al. 2003). La surexpression des gènes CBF1/*DREB1b* ou CBF3/*DREB1a* chez des plantes transgéniques d'*Arabidopsis* induit les gènes COR qui contiennent le CRT/DRE sans stimulation par la basse température. De plus, cette surexpression augmente la tolérance au gel en comparaison aux plantes contrôles (non acclimatées) (Thomashow, 2001).

Le domaine AP2 de la protéine CBF et la signature CBF avoisinante de ce domaine donnent la spécificité pour le ciblage des CBFs sur le motif de liaison à l'ADN, le CRT/DRE. Les analyses sur la protéine CBF d'*Arabidopsis*, nous montrent que les premiers 115 acides aminés du côté N-terminal de CBF1 sont responsable du ciblage vers les gènes contenant un CRT/DRE et les 98 acides aminés du côté C-terminal sont responsable de l'activité de transactivation (Wang et al. 2005). Pour étudier plus en détail l'interaction entre CBF et les autres protéines accessoires à la transcription des gènes cibles, Wang et ses collègues (2005) ont fait une dissection moléculaire de la portion C-terminale de CBF1. Ce domaine possède plusieurs

noyaux hydrophobiques HC2, HC3, HC4, HC5 et HC6 (hydrophobic clusters) qui sont bordés par de courtes régions composées de résidus favorisant la formation de boucles, et ceux-ci sont en combinaison avec les résidus de proline. Chez les mutants d'*Arabidopsis*, des résultats très similaires ont été obtenus pour différentes mutations de HC. Il semble que la région C-terminale est remarquablement tolérante à la substitution d'acides aminés. Les analyses de mutants et la modélisation indiquent que le domaine d'activation de CBF contribue une redondance fonctionnelle de transactivation. Chaque motif apporte une contribution significative vers la capacité totale de transactivation de CBF1, mais l'élimination d'un motif unique n'est pas suffisante pour abolir l'activité. Mais lorsque deux motifs ou plus sont mutés et combinés dans une seule construction, le potentiel de transactivation est compromis. Malgré la forte conservation des résidus dans les noyaux hydrophobiques entre les plantes, des mutations ponctuelles sur un acide aminé semblent avoir un effet négligeable sur l'activité de transactivation (dans HC4, des substitutions de M183, M187, L188 et L189 par l'alanine). La comparaison des activités de différents mutants a dévoilé que deux résidus adjacents de HC4, L179 et L180, représentent la majeure partie des propriétés d'activation du motif HC4. Chez les monocotylédones, ce sont plutôt des résidus Tyr qui se trouvent à ces positions, et également chez la plupart des monocotylédones, il manque une Pro pour séparer les motifs HC3 de HC4. Il semble que la présence d'un résidu de Trp est importante pour la transactivation par HC5, une substitution de Trp par une alanine montre une diminution de la capacité de transactivation de HC5 de CBF1 chez *Arabidopsis*.

Pour comprendre le mécanisme par lequel les gènes *CBF* améliorent la tolérance au gel, McKhann et ses collègues (2008) ont analysé la relation entre les polymorphismes dans ces gènes et leur expression ainsi que l'expression de leurs gènes cibles, *COR*, chez différentes origines géographiques d'*Arabidopsis*. Les gènes *CBF* sont fortement polymorphiques dans leurs promoteurs ainsi que dans les régions codantes. Une étude portant sur huit variétés d'*Arabidopsis* qui sont résistantes ou sensibles au froid a révélé que deux variétés d'*Arabidopsis* sensibles au froid montrent des insertions dans le promoteur de *CBF1* (McKhann et al. 2008). Les insertions semblent être identiques, et il n'y a aucune homologie avec d'autres séquences d'*Arabidopsis*. Ces



deux variétés montrent une expression différente des *CBF*, qui ne semble pas être à cause des insertions, mais plutôt à cause de l'existence d'un certain nombre de polymorphismes dans le promoteur entre ces deux variétés. Chez quatre cultivars sensibles au froid, les protéines *CBF* sont plus polymorphes. L'étude des trois gènes *CBF* (1, 2 et 3) chez quatre cultivars sensibles d'*Arabidopsis* a dévoilé que parmi les 21 polymorphismes affectant les acides aminés, neuf se trouvent dans les quatre cultivars sensibles au froid. Quelques substitutions non conservatrices d'acides aminés pourraient causer des changements dans la structure des protéines *CBF* et par la suite l'expression de gènes *COR* (McKhann et al. 2008).

### 1.7.1 GÈNES *CBF* CHEZ L'ORGE

Chez l'orge, il y a 20 gènes *CBF* qui représentent 14 gènes *HvCBF* différents. L'analyse phylogénétique des *CBFs* chez l'orge révèle trois sous-groupes majeurs. Tous les *CBFs* chez les monocotylédones se retrouvent dans un de ces trois sous-groupes (Skinner et al. 2005). Le sous-groupe *HvCBF1* est ancestral de *HvCBF3* et *HvCBF4*, et est composé de 4 gènes *HvCBF5*, *HvCBF7*, *HvCBF1* et *HvCBF11*. Les *CBFs* des monocotylédones partagent quelques caractéristiques contextuelles générales avec les *CBFs* des dicotylédones. Comme chez les dicotylédones, les séquences génomiques des *CBF* des monocotylédones ne possèdent pas d'intron. La structure de la protéine est aussi conservée et contient une séquence leader (N-terminale) d'environ 15 à 40 acides aminés, un domaine de liaison à l'ADN AP2 flanqué de chaque côté du motif spécifique de la sous-famille *CBF*, et un domaine C-terminal acide qui agirait comme région d'activation de la transcription.

L'expression de gènes *CBF* est typiquement induite par les stress abiotiques. Skinner et ses collègues (2005) ont évalué la réponse au froid des *HvCBFs* chez le cultivar Dicktoo, un génotype résistant au gel, et Morex, un génotype sensible au gel. Un traitement au froid induit fortement les trois gènes du sous-groupe *HvCBF4* (*HvCBF2*, *HvCBF4* et *HvCBF9*) entre 1 et 4 h d'exposition, avec le maximum d'expression survenant entre 4 et 8 h. Le niveau de réponse est

différent selon le génotype. *HvCBF2* et *HvCBF9* sont induits plus fortement chez le cultivar Dicktoo tandis que *HvCBF4* est induit à un niveau plus élevé chez le cultivar Morex.

Basé sur les expériences de Skinner et ses collègues (2005) chacun des trois membres du sous-groupe *HvCBF3* (*HvCBF3*, 6, et 8) montre un profil d'expression distinct. *HvCBF3* est faiblement exprimé et la transcription est à peine détectable comme une induction légère après 1 h d'exposition au froid chez le cultivar Dicktoo. L'expression de *HvCBF6* est plus forte que celle de *HvCBF3* et elle est détectable après 8 h de traitement au froid, puis atteint un pic après 24 h, et ce niveau est maintenu. *HvCBF8* est exprimé plus fortement en absence de stress et est régulé négativement par le stress salin, la sécheresse et la basse température. Les deux gènes du sous-groupe *HvCBF1* montrent également des réponses distinctes. *HvCBF1* est rapidement induit par le froid chez les deux génotypes, mais de façon transitoire. *HvCBF7* est induit par le froid et contrairement à l'expression de *HvCBF1* qui est plus forte chez le cultivar Dicktoo, l'expression de *HvCBF7* est plus forte chez le cultivar Morex.

Une des caractéristiques importantes des CBF est leur capacité à se lier au motif CRT, l'élément impliqué dans la régulation par le froid de plusieurs gènes *COR*. Pour vérifier si les CBFs de chaque sous-groupe phylogénétique chez l'orge sont compétents pour lier physiquement le motif CRT, Skinner et ses collègues (2005) ont évalué leur liaison avec des segments d'un promoteur d'*Arabidopsis* (*COR15a*) et deux promoteurs d'orge (*cor14b*, *Dhn5*). Les résultats montrent que les CBFs ont différentes affinités pour les trois motifs CRTs. Ces différentes affinités signifient que le contexte des séquences avoisinantes au motif CRT influence la liaison des HvCBF. Ces résultats démontrent que les CBFs chez l'orge montrent le critère fonctionnel de liaison spécifique à un motif CRT (Skinner et al. 2005).

Le motif CRT/DRE (CCGAC) est responsable de l'expression de plusieurs gènes par la voie indépendante de l'ABA chez *Arabidopsis* (Jiang et al. 1996). Chez l'orge, l'élément cis CCGAAA a été impliqué dans le processus de réponse au froid (Dunn et al. 1998). Toutefois, nombreux gènes inductibles par le froid chez l'orge, comme *HVA1* (Straub et al. 1994) et *blt4.6*

(White et al. 1994), ne contiennent pas le motif CCGAAA. Xue (2002) a observé que les promoteurs synthétiques contenant un motif G/ACCGAC ou un motif CCGAAA ont été activés par HvCBF1. Cependant, l'analyse comparative des niveaux d'induction par HvCBF1 a révélé que le motif CCGAAA provenant du promoteur de *blt4.9* a été beaucoup moins efficace que le motif GCCGAC. L'expérience de substitution de base a révélé que seulement la séquence CGAC du motif GCCGAC est importante pour l'interaction avec HvCBF1. Il semble que la base A à la position C<sub>6</sub> du motif (G<sub>1</sub>C<sub>2</sub>C<sub>3</sub>G<sub>4</sub>A<sub>5</sub>C<sub>6</sub>) n'est pas une base efficace pour HvCBF1. Cela explique pourquoi le niveau d'expression du gène *blt4.9* (contient un motif CCGAAA) activé par HvCBF1 a été extrêmement faible. L'activité des promoteurs contenant des motifs mutants, ACCGAC et GTCGAC, était de 36% et 75% de celle contenant le motif GCCGAC, respectivement. Le site préféré de liaison de la protéine HvCBF2 sur le gène *COR* est le motif (G/a) (T/c) CGAC. La protéine de HvCBF2 est un facteur de transcription qui se lie préférentiellement au motif GTCGAC du gène *HVA1*. La protéine HvCBF2 est seulement active au froid (Xue, 2003). Il semble que les fonctions biologiques de HvCBF1 et HvCBF2 sont similaires. Les deux protéines se lient au promoteur de *HVA1*. La séquence préférée pour la liaison de HvCBF2 est très proche de celle de HvCBF1 (Xue, 2003). La composition des bases entourant le motif influence fortement l'affinité de liaison de ces CBFs. Les sites préférés des deux HvCBFs ne sont pas présents sur le promoteur de *HVA1*, le gène induit par les basses températures et la sécheresse chez l'orge. Le promoteur de *HVA1* contient des motifs CRT similaires ATGCCGACGC et CCATCGACAT. Toutefois, l'interaction d'un facteur de transcription avec un motif de gène cible ne nécessite pas un appariement parfait (Seki et al. 2001). Comme le gène rapporteur sous le contrôle du promoteur de *HVA1* peut être activé par HvCBF2 dans la cellule de feuille traitée par le froid chez l'orge à un niveau similaire à HvCBF1 (Xue, 2003), alors HvCBF2 est potentiellement impliqué dans l'induction de *HVA1* chez l'orge au froid.

### 1.7.2 GÈNES *CBF* CHEZ LE BLÉ

Le blé est un modèle pour étudier les gènes *CBF* car les trois génomes du blé (hexaploïde) contiennent des gènes qui sont différents (Gill et al. 2004). Badawi et ses collègues (2007) ont identifié 37 gènes *CBF* différents qui représentent 15 groupes de gènes (1 à 3 copies par groupe) chez le blé hexaploïde. Une analyse des *CBF* présent chez les céréales suggère que *T. aestivum*, peut avoir près de 25 groupes de gènes *CBF* différents. L'analyse phylogénique de ces gènes montre que les *CBFs* se divisent en 10 groupes monophylétiques. Badawi et ses collègues (2007) ont proposé la nomenclature suivante pour les sous-groupes *CBF* : *CBFI*, II, IIIa, IIIb, IIIc, IIId, IVa, IVb, IVc, IVd. En se basant sur les analyses structurales des différentes protéines *CBF*, ils proposent que chaque sous-groupe ait des fonctions spécifiques et différentes des autres sous-groupes. Les groupes IIIc, IIId, IVa, IVb, IVc et IVd se trouvent spécifiquement chez les Pooideae.

Les gènes des groupes *CBFI* (*TaCBFI-A11*) et *CBFIIIc* (*CBFIIIc-3.1*, *CBFIIIc-3.2*) et le gène *CBFIVb-21.1* ne sont pas induits par le froid, tandis que ceux du groupe *TaCBFII* s'expriment de façon constitutive. Tous les autres *TaCBFs* montrent une induction transitoire rapide après exposition au froid, ils sont peu ou pas exprimés avant le traitement au froid et leur expression est au maximum après 4 à 6 heures de traitement. Le niveau d'expression revient à la normale après 1-7 jours au froid. La comparaison des cinétiques d'induction des *CBF* chez les cultivars d'hiver et printemps dévoile qu'elles sont similaires (Badawi et al. 2007). Toutefois, l'expression des gènes des groupes *CBFIIId*, *CBFIVa*, *CBFIVb*, *CBFIVc* et *CBFIVd* (sauf *TaCBFIVd-D9*) montre une augmentation plus élevée chez le cultivar d'hiver que chez le cultivar de printemps. Cette étude quantitative révèle que huit gènes *CBFs* (*IIId-B12*, *IIId-15.2*, *IIId-A19*, *IVa-A2*, *IVb-D20*, *IVb-D21*, *IVd-B14* et *IVd-B22*) montrent une augmentation d'expression de cinq fois ou plus chez le blé d'hiver lors d'une exposition au froid. Cette propriété d'expression est un caractère hérité qui peut augmenter la capacité de tolérance à la basse température chez les cultivars d'hiver (Badawi et al. 2007). De plus, les gènes *CBFs* des



groupes IIIId, IVa, IVb, IVc et IVd ont montré une régulation par le cycle circadien à des conditions normale de croissance (20°C) (Badawi et al. 2007).

Badawi et ses collègues (2008) ont étudié ICE1, un des inducteurs d'expression de *CBF* chez le blé traité au froid. Ils ont isolé deux gènes *TaICE41* et *TaICE87* qui codent respectivement pour une protéine de 381 acides aminés (39,5 kDa) et pour une protéine de 443 acides aminés (46,5 kDa). *TaICE41* et *TaICE87* sont identique à 46% alors qu'elles partagent 50 et 47% d'identité avec *AtICE1* chez *Arabidopsis*, respectivement. En plus, *TaICE41* et *TaICE87* se lient à différents éléments MYC sur le promoteur de *TaCBFIVd-B9* (Badawi et al. 2008). Les différentes affinités de *TaICE41* et *TaICE87* pour les variantes MYC suggèrent que la spécificité de liaison ICE peut être impliquée dans l'expression différentielle des gènes *CBF* du blé. En outre, l'analyse des éléments MYC démontre qu'une particulière variante est présent dans les CBFs des groupes IV chez le blé, des groupes associés à la tolérance au gel. Une hypothèse plausible pour expliquer les différences d'expression entre les cultivars hiver et les cultivars de printemps pourrait être que les gènes *CBF* possèdent des différences moléculaires au niveau de ces éléments dans leur promoteur qui permettent d'induire des expressions plus élevées de certains gènes *CBF* chez les cultivars d'hiver.

## 1.8 VERNALISATION ET SA RELATION AVEC LA RÉSISTANCE AU GEL

Dans les régions tempérées, la basse température constitue un obstacle majeur qui régule le temps de floraison et de nombreuses transitions développementales, comme la germination, la dormance des bourgeons et leur éclatement (Henderson et al. 2003). La vernalisation se réfère à une période de froid subie par la plante ou la graine en germination, et nécessaire pour la faire passer du stade végétatif au stade reproductif, donc de fleurir. Les variétés de printemps ne répondent pas à la vernalisation et donc fleurissent rapidement tandis que les variétés d'hiver ont besoin de vernalisation pour accélérer la floraison et achever leur cycle de vie. Chez les céréales, la vernalisation et l'acclimatation au froid sont deux processus reliés entre eux. Plusieurs études

chez l'orge ont rapporté que l'expression de gènes *CBFs* et plus tard des gènes *COR*, est réduite après l'induction de *Vrn-H1* à la réponse de vernalisation, ce qui suggère une interaction entre les réseaux de régulation génétique de la vernalisation et d'acclimatation au froid (Stockinger et al. 2007). Le blé a trois loci majeurs de vernalisation, *Vrn-A1*, *Vrn-B1*, et *Vrn-D1*, qui sont tous liés aux loci *Fr-1* (Galiba et al. 2009; Snape et al. 1998; Iwaki et al. 2002; Toth et al. 2003). L'analyse génétique et les études sur l'expression des gènes *Vrn-1* chez le blé ont montré que l'accumulation de transcrits de ce gène est associé à la réponse de vernalisation et la transition de la phase végétative à la phase reproductive (Danyluk et al. 2003; Murai et al. 2003; Yan et al. 2003). Chez le blé et l'orge, *VRN2* est un autre facteur de transcription de type doigt de zinc qui agit comme un régulateur négatif dominant de la période de floraison dans la voie de la vernalisation (Karsai et al. 2005; Yan et al. 2004). La réduction du niveau d'expression de *VRN2* par RNAi accélère la floraison du blé hexaploïde (Yan et al. 2004). *VRN3* régule positivement la floraison sous condition de longue journée, par régulation de *VRN1*, le gène de l'identité du méristème. L'effet de l'augmentation de la transcription de *VRN1* est la régulation négative de *VRN2* (Trevaskis et al. 2006; Dubcovsky et al. 2005). La présence et l'expression de l'allèle dominant du gène *Vrn1* chez le blé de printemps montre une relation inverse avec le maximum de tolérance au gel (Fu et al. 2005). Le gène *Vrn* contrôle non seulement la transition du stade végétatif au stade reproductif, mais aussi affecte l'expression des gènes induits à la basse température qui sont associés à l'acquisition de la tolérance au gel (Danyluk et al. 2003).

## 1.9 RÔLES POSSIBLES D'UNE AMPLIFICATION DU CONTENU GÉNIQUE

En utilisant une importante collection de lignées de plantes à fleurs, y compris les graminées, une étude a révélé que les plantes qui ont subi une duplication de génome entier ont été plus susceptibles de survivre à la période d'extinction massive de la limite Crétacé-Tertiaire (K-T) (Fawcett et al. 2009). D'autre part, l'augmentation de copie de gènes comme adaptation génétique ne se limite pas à la duplication de génome entier. Une expansion sélective de gènes spécifiques peut également jouer un rôle important dans l'évolution adaptative (Kondrashov et al.

2006; Chen et al. 2008; Zou et al. 2009). Sur ce point, Sandve et Fjellheim (2010) ont proposé un modèle d'évolution de la tolérance au gel chez les *Pooideae* où une forte pression de sélection sur un de leurs ancêtres pendant la période Éocène oligocène a conduit à l'évolution adaptative de la tolérance au gel grâce à une expansion des familles de gènes *CBFIIIc/d* et *CBFIV*. Le modèle d'évolution pour la conservation de gènes dupliqués prévoit plusieurs mécanismes dans lesquels un événement de duplication génique pourrait être adaptatif. Tout d'abord, un avantage immédiat d'un événement de duplication génique est soit un effet de l'augmentation de la dose génique (Ferrari et Georgiou, 1991) ou la redondance fonctionnelle (Gu, 2003; Moore et al. 2005). L'effet d'une dose plus élevée des gènes *CBFIIIc/d* et *CBFIV* pourrait conduire à une expression accrue des facteurs de transcription sous des conditions de stress abiotiques. Cela pourrait nous amener (i) à une réponse plus rapide au stress dû à une accumulation plus rapide des produits des gènes, ou (ii) une réponse plus longue et plus forte au stress au froid, du à la persistance des ARNm de gènes CBF dans la cellule pour un temps plus long.

Un autre modèle expliquant la conservation de gènes dupliqués est la spécialisation des gènes, sous-fonctionnalisation et néo-fonctionnalisation (Ohno, 1970; Gibson et Goldberg, 2009; Hahn, 2009). Ce modèle suppose une sélection positive sur des mutations post-duplication plutôt que la conséquence d'adaptation de l'événement de duplication elle-même (Hahn, 2009). Ceci amènerait une évolution de lignées spécifiques des gènes et de leur expression qui serait impliquée dans l'évolution adaptative (Ferea et al. 1999; Rifkin et al. 2003; Nuzhdin et al. 2004). Les études suggèrent que la spécialisation, sous- ou néo-fonctionnalisation des gènes *CBFIIIc/d* et *CBFIV* peut être le résultat des changements favorables sous les stress abiotiques conduisant à des nouveaux modèles d'expression génétique et l'amélioration de la réponse au froid.

#### 1.10 QTL (QUANTITATIVE TRAIT LOCUS)

Un locus de caractères quantitatifs (LCQ ou QTL pour quantitative trait locus) est une région plus ou moins grande d'ADN qui est étroitement associée à un caractère quantitatif, c'est-



à-dire une région chromosomique où sont localisés un ou plusieurs gènes à l'origine du caractère en question. Un QTL nous aide à relier deux types de données, l'information phénotypique et les données génotypiques. Une analyse QTL permet aux chercheurs dans des domaines aussi variés que l'agriculture, l'évolution et la médecine de relier certains phénotypes complexes à des régions spécifiques des chromosomes. Le but de ce processus consiste à identifier l'action, l'interaction, le nombre et l'emplacement précis de ces régions (Miles et Wayne, 2008). Par exemple, chez l'humain, les chercheurs ont trouvé plusieurs mutations à la base de problèmes de santé. Les facteurs génétiques jouent un rôle important dans le cancer du sein et des ovaires, mais moins de 10% sont attribuables à l'héritage de mutations dans un seul gène (Wang et al. 2001).

Les Single-Nucleotide Polymorphisms (SNP) sont, en génétique, des variations (ou polymorphismes) d'une seule paire de bases du génome, entre individus d'une même espèce. Le SNP peut survenir sur le promoteur ou dans la partie codante du gène. Le SNP dans une séquence codante ne change pas nécessairement l'acide aminé de la protéine produite par le gène, en raison de la dégénérescence du code génétique. Un SNP qui produit la même séquence polypeptidique est appelé un polymorphisme synonyme (aussi mutation silencieuse). Chez les humains, 99,9% des séquences d'ADN individuel sont identiques à celui d'une autre personne. Les SNP représentent 80% de l'ensemble des variations génétiques humaines. Ces variations sont très fréquentes (environ 1 par 1000 paires de bases dans le génome humain). La présence d'un SNP (G/C 135) sur le gène *RAD51*, semble avoir un effet sur l'apparition des cancers du sein et de l'ovaire (Wang et al. 2001). Chez les humains, un SNP est à la base du problème de la drépanocytose (également appelée anémie à cellules falciformes) qui est une maladie héréditaire qui se caractérise par l'altération de l'hémoglobine, la protéine assurant le transport de l'oxygène dans le sang. La drépanocytose n'est pas une maladie très rare. Elle est particulièrement fréquente dans la population d'origine africaine subsaharienne, des Antilles, de l'Inde, du Moyen-Orient et du bassin méditerranéen. On estime que 50 millions d'individus en sont atteints dans le monde (Arnal et Girot, 2002). L'allèle S, est un allèle anormal du gène régissant la structure de la chaîne

bêta de l'hémoglobine. Dans cet allèle, un acide glutamique en position 6 est remplacé par une valine (Bardakjian-Michau, 2008).

#### 1.10.1 IDENTIFICATION DE QTL CHEZ LES CÉRÉALES SUR DES GÈNES ASSOCIÉS À D'AUTRES STRESS QUE LE FROID

##### 1.10.1.1 QTL chez les céréales

La présence de plusieurs QTL a été identifiée chez les céréales, telles que ceux associés à la tolérance au stress salin, la tolérance au gel, la tolérance au stress d'aluminium (Al) et plusieurs autres. Ici, nous présentons quelques exemples de QTL retrouvés chez les céréales.

Le principal obstacle à la croissance des plantes poussant sur les sols acidifiés est la toxicité par l'aluminium (Al) qui cause une inhibition de la croissance des racines. Il existe une variation significative dans la capacité des génotypes au sein de certaines espèces résistantes, d'exclure l'Al de leurs tissus, ou de tolérer l'Al par détoxification ou la séquestration dans certains organites et tissus (Kochian, 1995; Matsumoto, 2000; Kochian et al. 2004). Dans l'étude de Sasaki et ses collègues (2006) sur le blé, 69 lignées ont été examinées pour leur variabilité dans le gène de résistance au stress d'Al (*ALMT1*) et l'analyse a montré que contrairement aux séquences relativement conservées de la région codante du gène, il y avait de la variabilité dans le premier 1 kpb en amont du codon d'initiation de ce gène. Les régions en amont de nombreux gènes dont leur expression est induite par les stress abiotiques (sécheresse et le froid) contiennent des éléments *cis* qui sont impliqués dans la réponse au stress (Yamaguchi-Shinozaki et Shinozaki, 2005). Une recherche sur la séquence du promoteur des gènes de résistance à l'Al a permis d'identifier une répétition du promoteur chez les cultivars résistants ce qui augmente l'expression de ce gène (Sasaki et al. 2006).

Un autre exemple est l'existence d'une variation génétique pour la tolérance à la toxicité du Bore (B) chez l'orge. Le contenu en B dans les sols arides de l'Australie et de nombreuses autres parties du monde peut causer des pertes de rendement chez des variétés d'orge sensibles (Cartwright et al. 1984). Oertli et Roth (1969) ont suggéré que la tolérance au B est causée par l'accumulation réduite de B. Cela a été confirmé par Nable (1988) qui a trouvé que les blés et orges résistants accumulent moins de B dans leurs racines et pousses. Huang et Graham (1990) et Choi (2004) ont montré que des changements dans l'architecture des racines sont associés avec la tolérance au B chez le blé et l'orge. En utilisant différents croisements entre un cultivar B-tolérant et un cultivar B-sensible, plusieurs QTL ont été identifiés avec des effets détectables sur l'accumulation de B dans la pousse, sur l'architecture des racines, la production de matière sèche et l'expression des symptômes chez les feuilles (Jefferies et al. 1999; 2001). Chez l'orge, les loci sur les chromosomes 4H et 6H agissent additivement et contribuent pour 64% et 31% de la réduction de l'accumulation de bore, respectivement. Récemment, *Bot1* a été identifié comme le gène de tolérance (Reid, 2007; Sutton et al. 2007). Chez le cultivar résistant, il existe en tandem de quatre exemplaires sur le chromosome 4H (Sutton et al. 2007). L'abondance des transcrits de *Bot1* indique que ce gène est plus abondamment exprimé chez le génotype tolérant (Hassan et al. 2010).

#### 1.10.1.2 Identification des QTLs de gènes CBFs chez le blé et l'orge

Chez le blé diploïde (*T. monococcum*), onze gènes *CBF* ont été cartographiés sur une région de 0,8 cM sur le chromosome 5, région coïncidant avec le QTL *Frost resistance-2* (*Fr-A<sup>m</sup>2*) (Miller et al. 2006). Ce locus est associé à une plus grande tolérance au gel et à une expression différentielle du gène *COR14b* par le froid chez le cultivar plus tolérant (Vágújfalvi et al. 2000; 2003). Actuellement, il n'est pas encore connu quel(s) gène *CBF(s)* au sein de ce cluster est/sont responsable(s) des différences observées dans la tolérance au gel (Vágújfalvi et al. 2005; Stockinger et al. 2007). Une stratégie pour trouver le ou les gènes *CBF* de ce locus est de générer des populations de cartographie à haute densité, et d'identifier les événements de recombinaison



entre les gènes *CBF* dans ce locus. Knox et ses collègues (2008) ont testé des lignées montrant des événements de recombinaison pour la tolérance au gel et pour l'expression différentielle de *COR14b*. Leurs différentes analyses ont indiqué que l'effet majeur était lié au centre du cluster de *CBF*, y compris, *TmCBF12*, *TmCBF14* et *TmCBF15*. La différence la plus évidente sur ces séquences de gènes candidats est une délétion dans le domaine AP2 de *TmCBF12* chez le cultivar sensible au gel. Le domaine AP2 des protéines CBF se lie au motif CRT/DRE (CCGAC) qui est présent dans les promoteurs de nombreux gènes induits au froid (Baker et al. 1994; Yamaguchi-Shinozaki et Shinozaki, 1994; Stockinger et al. 1997; Gilmour et al. 1998; Liu et al. 1998; Thomashow, 1999).

La structure du domaine de liaison à l'ADN, le domaine AP2 des protéines CBF, est composée de trois feuillets  $\beta$  ( $\beta 1$ -3) antiparallèles, et d'une hélice  $\alpha$  presque parallèle aux feuillets  $\beta$  (Allen et al. 1998). La délétion de 5 acides aminés chez *TmCBF12* du cultivar sensible au froid est située entre les feuillets  $\beta 1$  et  $\beta 2$ , dans une région connue pour se lier au motif CCGAC sur le promoteur des gènes *COR* (Allen et al. 1998; Hao et al. 2002). Les résultats obtenus par gel de retardement ont confirmé que cette délétion chez *CBF12* élimine la capacité de cette protéine à se lier au promoteur d'un gène *COR*. Par conséquent, il est probable que cette protéine ne peut plus moduler l'expression des gènes *COR*, et ainsi altère la capacité de la plante à s'acclimater au froid et à développer la tolérance au gel.

La comparaison de séquences conservées dans le domaine AP2 de protéines CBF montre que leurs moitiés N-terminale sont presque identiques. Cette région semble coïncider avec la spécificité de liaison pour la boîte GCC, ce qui implique que la moitié N-terminale est responsable de la reconnaissance spécifique de la boîte GCC (Hao et al. 1998). Allen et ses collègues (1998) ont mentionné que la structure du domaine AP2 de la protéine ERF1 chez *Arabidopsis* est stabilisée par de nombreux contacts hydrophobiques entre les chaînes latérales des résidus Tyr146, Val149, Phe157, Ala159, Ile161, Val171, Leu173, Phe176, Ala179, Ala182, Ala183, Ala185, Tyr186, Ala189, Ala190, Ala198, Leu200 et Phe202. La géométrie de l'hélice- $\alpha$  vis-à-vis les feuillets- $\beta$  semble être déterminée par l'interaction de nombreux résidus Ala dans

l'hélice- $\alpha$  avec les plus grands résidus hydrophobes dans les feuillets- $\beta$ , en particulier, Phe157, Phe176, Val171 et Ile161, qui pince l'hélice aux quatre coins. Les groupes guanidyl de quatre résidus arginine ont été identifiés comme ayant des contacts avec cinq bases guanine par des liaisons hydrogènes (avec la boîte GCC), Arg 150 avec G20, Arg 152 avec G5 et G21, Arg 162 avec G17 et Arg 170 avec G8. Les chaînes latérales de résidus Arg 150, Arg 170 et Arg 162 ont des interactions hydrophobes avec des bases pyrimidiques C19, C7 et T8, respectivement. Les cycles aromatiques du Trp154 établissent des interactions hydrophobes avec les bases T3 et A4 et ceux du Trp172 avec, G5 et C6 sur la chaîne d'ADN. Ces interactions couvrent les six paires de bases dans la séquence conservée AGCCGCC et alors servent à la reconnaissance spécifique de la boîte GCC par ce facteur. Tous les résidus Arg et Trp qui contactent les bases (sauf Arg 152), ont, en plus, des liaisons hydrophiles avec le sucre ou des liaisons ioniques avec des phosphates de l'ADN. De plus, certains autres résidus (Lys156, Thr175, Arg147, Tyr186 et Gly148) ont des liaisons avec le sucre/phosphate de l'ADN (Allen et al. 1998).

De plus les CBF possèdent une séquence signature hautement conservée en amont du domaine AP2 suggérant qu'elle a une fonction importante (Canella et al. 2010). Canella et ses collègues (2010) ont effectué des analyses afin de déterminer s'il y a des acides aminés spécifiques dans cette signature qui sont importants pour la liaison à l'ADN. Les résultats indiquent que les 16 acides aminés en amont du domaine AP2 ne peuvent pas avoir un rôle direct dans la liaison de l'ADN, mais plutôt sont requis pour le domaine AP2/ERF pour maintenir une conformation active et en plus, les résidus Arg37 et Phe40 ont été critiques pour la liaison de l'ADN et les substitutions de ces acides aminés ont toutes fortement aboli la liaison aux éléments régulateurs CRT/DRE.

Le blé d'hiver est plus résistant au froid que son homologue, le blé de printemps (Limin et Fowler, 2006; Wilen et al. 1996). L'expression des gènes *CBFs* et la durée de l'expression influence la capacité finale de la résistance au gel (Mahfoozi et al. 2000; Fowler et al. 2001). Stockinger et ses collègues (2007) ont trouvé qu'il n'y a pas de différence qualitative entre l'expression de gènes *CBFs* de blé d'hiver et de blé de printemps, mais que ces gènes sont

exprimés chez le blé sensible au froid à des niveaux inférieurs. L'expression accrue d'un gène peut être tout simplement causée par une augmentation de son nombre de copies (Knox et al. 2010). Alors, Knox et ses collègues (2010) ont vérifié s'il y a des différences structurelles entre les gènes *CBF* chez l'orge résistant et l'orge sensible au froid qui pourrait expliquer la différence phénotypique. Les résultats montrent que le cultivar de printemps possède un seul gène *CBF2* sous une forme chimère *CBF2B/A* alors que chez le cultivar d'hiver, ce gène existe en deux copies *CBF2B* et *CBF2A*. Les résultats également montrent que le cultivar résistant au froid contient au moins deux copies du gène *CBF4B* en tandem à la suite d'une duplication de la région génomique comprenant *CBF4B*. La deuxième différence est sur le gène *CBF13*. Le gène *CBF13* chez le cultivar résistant est une séquence codante intacte, alors que chez le cultivar sensible c'est un pseudogène.

Chez l'orge, 20 gènes *CBF* ont été identifiés (Skinner et al. 2006), dont 11 sur une région génomique du chromosome 5H, déjà identifiée comme locus contenant le QTL (*Fr-H2*) (Skinner et al. 2006; Francia et al. 2007). L'analyse de la structure génomique a révélé une organisation non conservée dans le locus *Fr-H2* de l'orge: les cultivars Nure et Dicktoo (deux cultivars d'hiver) se partagent la même organisation génomique, alors que Morex et Tremois (deux cultivars de printemps) sont caractérisés par une autre structure génomique (Francia et al. 2007). Comme la tolérance au gel est une caractéristique des plantes de climats tempérés, les gènes *CBF* devraient être hautement conservés dans la tribu *Triticeae*. Fricano et ses collègues (2009) ont travaillé sur les variations alléliques de quatre gènes *CBF* chez l'orge (*HvCBF3*, *HvCBF6*, *HvCBF9* et *HvCBF14*) dans un groupe de cultivars européens, chez des variétés locales et *H. spontaneum*. Le but était de trouver une association entre des variantes génétiques chez plusieurs gènes *CBF* et la co-ségrégation avec le QTL de résistance au gel. Il est généralement admis que les plantes cultivées contiennent moins de variation génétique que leurs ancêtres sauvages (Doebley et al. 2006). En plus, le processus de domestication peut impliquer une réduction de variation génétique causée par la reproduction répétée d'individus supérieurs (Doebley et al. 2006). L'analyse des ressources génétiques *CBF* présentes chez des espèces cultivées européennes a révélé que les séquences de *HvCBF9* sont presque identiques chez les variétés



cultivées, et absentes dans certaines accessions d'orge. L'utilisation d'amorces additionnelles n'a pas permis de détecter *HvCBF9* dans certaines lignées. Les variants de *HvCBF9* n'étaient pas associés à un niveau plus élevé de tolérance au gel, suggérant que ce gène n'a pas d'effet majeur sur ce trait. Comme la sélection ciblée sur un locus individuel réduit la diversité génétique au sein de et autour du locus sélectionné (un phénomène connu sous le nom de genetic hitchhiking; Smith et Haigh, 1974; Palaisa et al. 2004), la forte réduction des allèles de *HvCBF9* chez les variétés cultivées auraient pu être générées par la sélection ciblée aux gènes *CBF*. Une situation similaire est valable pour *HvCBF3* qui n'a pas été détectée dans un tiers des cultivars européens. L'analyse d'association entre variant de gènes *CBF* et la tolérance au gel montrent que deux substitutions nucléotidiques de *HvCBF14* sont statistiquement associées à la tolérance au gel. Les deux polymorphismes sont dans la région 3'-non codante de *HvCBF14* et l'analyse bioinformatique indique que cette région ne contient pas d'éléments cis. Un certain nombre de lignées sauvages qui sont tolérantes au gel ont les mêmes variantes de *HvCBF14*. Il n'est cependant pas établi que ces polymorphismes sont responsables de la tolérance au gel.

Une analyse moléculaire de 12 gènes provenant de plusieurs loci a été effectuée à partir d'un ensemble de 102 accessions uniques d'orge (Cuesta-Marcos et al. 2010). Pour les gènes situés dans le locus *FR-H2*, ces auteurs ont trouvé 19 SNP et cinq haplotypes pour *HvCBF3*, 17 SNP et six haplotypes pour *HvCBF6*, et huit SNP et six haplotypes pour *HvCBF9*. Toutefois, aucune mention de l'effet de ces polymorphismes sur la fonction possible des CBF n'a été déterminée.

Le blé hexaploïde Chinese Spring (contenant le chromosome 5A de *Triticum spelta*), est insensible à la vernalisation, il contient l'allèle *Vrn1* trouvé au niveau du chromosome 5A. *Vrn-A1* est l'allèle déterminant dans le processus d'inhibition de la vernalisation (Galiba et al. 1995). La cartographie des chromosomes 5A et 5D de descendants issus du croisement entre Chinese Spring (*T. Spelta* 5A) et Chinese Spring (Cheyenne 5A) (vernalisation sensible) ou Chinese Spring et Chinese Spring (Cheyenne 5D) montre que les gènes *Vrn1* et *Fr1* (gène de tolérance au gel) sont des loci liés au niveau des chromosomes 5A et 5D. De plus, il existerait une relation

homeoallélique entre *Vrn-A1* et *Vrn-D1* (Snape et al. 2001). Toth et ses collègues (2003) ont montré que chez les descendants issus du croisement entre Chinese Spring et Chinese Spring (Cheyenne 5B), les loci *Vrn-B1* et *Fr-B1* sont liés aussi. *Fr-B1* contient un QTL et serait un gène orthologue aux autres gènes *Fr* situés sur les chromosomes 5A et 5D. Ces observations montrent que *Vrn-1A* n'est pas le seul responsable de la sensibilité de Chinese Spring au gel, supportant l'hypothèse que le locus avoisinant contient les *CBFs* soient impliqué.

## CHAPITRE II MATÉRIELS ET MÉTHODES

### 2.1 IDENTIFICATION DE NOUVEAUX GÈNES *CBF* DU BLÉ

Dans une étude précédente, Badawi et ses collègues (2007) ont isolé 15 groupes de gènes *CBF* chez le blé hexaploïde, et concluent que cette espèce pourrait avoir jusqu'à 25 groupes de gènes *CBFs*. Dans le but de vérifier cette possibilité, nous avons assemblé des séquences *CBF* en ARNm virtuel à l'aide de CAP3 (<http://pbil.univ-lyon1.fr/cap3.php>) (Huang et Madan, 1999). Les séquences assemblées incluent des séquences de *T. monococcum*, des séquences partielles obtenues préalablement dans notre laboratoire, et des gènes complets/incomplets disponibles dans les bases de données NR (non redondante) et EST (expressed sequence tag) du NCBI. Ce premier assemblage a été mis à jour au fur et à mesure que de nouveaux gènes *CBF* ont été séquencés. En utilisant ces assemblages d'ARNm virtuel, d'une à trois amorces spécifiques (GSP) du gène ont été conçues à l'aide de Primer 3 (<http://frodo.wi.mit.edu/primer3/>) (Rozen et Skaletsky, 2000) pour la direction sens (région 5' de l'ARNm) ou anti-sens (dans la région 3' de l'ARNm).

### 2.2 ISOLATION DES GÈNES *CBF* PAR PCR

Nous avons utilisé les amorces conçues (GSP) pour isoler les gènes *CBF* exprimés chez le blé à partir de 4 bibliothèques d'ADNc construites auparavant (Houde et al. 2006). Brièvement, ces bibliothèques d'ADNc ont été préparées à partir de *Triticum aestivum* L. cv Norstar. Les quatre bibliothèques (L3-L6) ont été préparées à partir de pool d'ARNm obtenue de: (L3), tissu de racine de blé contrôle, de blé acclimaté au froid et lors d'un stress salin; (L4), parties aériennes lors d'un stress de déshydratation; (L5), des tissus du collet au cours de leur vernalisation et différents stades de développement de l'épi et la formation de graines du blé; (L6), la partie aérienne du blé (collet et feuilles) après une courte exposition à la basse température sous la lumière et à

l'obscurité. Tous les ADNc synthétisés ont été clonés dans le même sens dans le vecteur pCMV.SPORT6 (Invitrogen) en utilisant l'adaptateur *SaII* (GTCGACCCACGCGTCCG) et l'amorce/adaptateur *NotI* (GCGGCCGCCCT<sub>15</sub>). Le premier mélange réactionnel pour la synthèse d'ADNc contenait du dCTP méthylé pour éviter le clivage interne des ADNc par l'enzyme de restriction *NotI* utilisé pour générer les extrémités cohésives pour le clonage. Pour la L6, le "GeneRacer" kit (Invitrogen) a été utilisé avant la synthèse du premier brin pour produire une librairie contenant une forte proportion (95%) d'ADNc pleine longueur. Pour chaque librairie, six millions de transformants ont été obtenus, amplifiés et congelés à -80°C avec du glycérol. Pour préparer les plasmides pour l'expérience de PCR, 40 µl d'un stock de la librairie d'ADNc (>40 × 10<sup>6</sup> clones) ont été inoculés en milieu LB (100 ml) additionné d'ampicilline et cultivés à 30°C pour 6-8 h. Les plasmides ont été isolés avec la trousse QIAprep Miniprep (QIAGEN) et évalués sur gel.

Afin d'améliorer la spécificité de la réaction PCR, les GSP ont été conçus avec une T<sub>m</sub> plus élevée que l'amorce universelle M13F 5'-AGATCCCAAGCTAGCAGTTTTCCAGTCACGA-3' et M13R 5'-GAGCGGATAACAATTTTCACACAGGAAACAGCTATGA-3' trouvé dans le vecteur pCMV.SPORT6. Des listes d'amorces GSP utilisées pour l'ajout de séquence aux gènes *CBFs* incomplets et pour l'isolement de gènes *CBF* nouveaux sont présentées respectivement dans les tableaux S1 et S2. L'amplification a été effectuée en utilisant un GSP (0,6 µM) et l'amorce correspondante universelle (0,3 µM), 10 ng d'ADN plasmidique à partir d'une librairie d'ADNc, 1× ou 2× de la solution enhancer, et l'ADN polymérase Pfx en suivant les recommandations du manufacturier (Invitrogen). Brièvement, le PCR débutait avec une diminution progressive de la température jusqu'à la température d'appariement (annealing) et s'est terminé avec 45 cycles de 94°C pendant 20 s, 64°C pendant 20 s et 68°C pendant 150 s. Les produits de PCR ont été analysés sur un gel et les réactions qui ont produit des fragments de taille attendue ont été choisies pour un sous-clonage en utilisant la trousse de clonage (Zero Blunt TOPO PCR kit; pCR®-Blunt II-TOPO) d'Invitrogen. Les produits de sous-clonage ont été transformés dans les bactéries compétentes *E. coli* (TOP10) et étalées sur pétri LB-agar-kanamycine et incubées



pendant 18 heures à 37°C. Quelques colonies ont été utilisées pour ensemer des tubes contenant du LB/kanamycine. L'ADN de plasmide a été isolé par la méthode utilisant de la terre de diatomée. Après une incubation à 37°C pendant 18 heures avec agitation, la suspension bactérienne a été transférée dans un tube Eppendorf et centrifugée à 14000g pendant 1min. Le culot a été récupéré et suspendu dans 100 µl de solution contenant 25mM TrisHCl, 10mM EDTA, 100µg/ml ARNase A, pH 8. Le mélange a été bien vortexé et 100 µl d'une solution de lyse (0.2 M NaOH contenant 1% SDS) a été ajouté et bien mélangé. Dans le même tube, il a été ajouté 400 µl de la solution de neutralisation/dénaturation (0.75 M potassium-acetate, 0.5 M acide acétique, 37.5 mM TrisHCl, 15 mM EDTA, 4.5 M guanidine HCl). Après 10 min à la température de la pièce, le mélange a été centrifugé à 14000g pendant 10 min et le surnageant a été transféré sur une colonne de terre de diatomée. Le liquide pénètre dans la colonne avec une centrifugation de 5000g pendant 1 min. Un lavage a été effectué avec 600µl de solution de lavage (8.3 mM Tris-HCl pH 7.5, 2.1 mM EDTA, 83 mM NaCl, 42% éthanol) et suivi par une centrifugation de 5000g pendant 1 min. Une deuxième centrifugation a été faite pour sécher la terre de diatomée. Les plasmides ont été récupérés en mettant 50 µl de tampon TE (10 mM Tris-HCl, pH 8.5, contenant 1 mM EDTA). Les plasmides contenant la séquence désirée ont été traités par l'enzyme de restriction *EcoRI*. Les résultats ont été évalués par électrophorèse sur gel d'agarose, puis les plasmides ont été envoyés pour séquençage au Centre d'Innovation Génome Québec et Université McGill. Les séquences représentant des nouveaux gènes *CBF* ont été entièrement séquencées. Les séquences qui montraient un chevauchement (100% d'homologie) ont été fusionnées pour produire la plus longue séquence du gène possible (Tableau 1). Au moins deux clones provenant de réactions d'amplification indépendante ont été séquencés afin de s'assurer qu'il n'y a pas d'erreurs introduites à cause des PCR.

### 2.3 ANALYSES PHYLOGÉNÉTIQUES ET BIOINFORMATIQUES

Les fichiers FASTA de séquences nucléotidiques codant soit pour le domaine AP2 / séquences signatures avoisinantes ou soit pour toute la protéine CBF, et utilisés dans les diverses

analyses phylogénétiques sont présentés dans les figures complémentaires S1 et S2. Les séquences ont été alignées en utilisant le programme ClustalW du Biology Workbench (<http://workbench.sdsc.edu/>), ou MEGA5 (Tamura et al. 2011) et l'alignement raffiné manuellement. Les régions contenant des alignements ambigus ou des GAP ont été enlevées avant l'analyse. Le logiciel MEGA5 a été utilisé pour les analyses phylogénétiques d'abord en trouvant le meilleur modèle de nucléotide pour nos données parmi ceux disponibles dans MEGA. Tant le critère d'information bayésien que le critère d'information d'Akaike ont toujours choisi les modèles T92+G et HKY+G comme ceux décrivant le mieux le patron de substitution. Ces modèles ont été utilisés dans l'analyse des nucléotides par maximum de vraisemblance avec le test de bootstrap (500 répliquats) et ont donné des résultats similaires. L'analyse de groupements hydrophobes (HCA) (Gaboriaud et al. 1987; Callebaut et al. 1997) a été effectuée à l'aide de l'interface basée sur le Web à l'adresse: ([http://www.bioserv.rpbs.jussieu.fr / RPBS / cgi-bin / Ressource.cgi? chzn\\_lg = une & chzn\\_rsrc = HCA](http://www.bioserv.rpbs.jussieu.fr/RPBS/cgi-bin/Ressource.cgi?chzn_lg=une&chzn_rsrc=HCA)). En bref, les séquences protéiques de *T. aestivum*, *T. monococcum*, *H. vulgare* et *S. cereale* sont affichées sur un filet de double hélice dans laquelle les acides aminés hydrophobes (V, I, L, F, M, Y, W) sont contournés. Les résidus hydrophobes séparés par quatre résidus non hydrophobes ou plus, ou une proline, sont placés dans des noyaux distincts. Le caractère hydrophobique de ces noyaux représente la face interne des structures secondaires régulières (hélices ou feuillettes). Un autre acide aminé a été choisi pour être mis en évidence dans cette étude: la proline qui confère le plus grand obstacle à la chaîne polypeptidique. Ces informations sur la structure secondaire ont été mises en évidence sur un alignement de protéines CBF.



Tableau 2.1 La nomenclature et les caractéristiques de nouveaux gènes chez le blé hexaploïde.

Nom du Gène	Numéro d'accension	Méthode d'isolation	ADNc/ ADN longueur (pb)	Cadre de lecture ouvert (CDS)	Queue poly A détectée	Acides aminés totaux
<i>TaCBFI-1.1</i>		1	1004	59-712	Oui	217
<i>TaCBFI-1.2</i>		1	1062	58-711	Oui	217
<i>TaCBFI-1.3</i>		1	1207	59-712	Oui	217
<i>TaCBFII-5.4</i>		1	1038	85-717	Oui	210
<i>TaCBFII-5.5</i>		1	1075	92-766	Oui	224 <sup>a</sup>
<i>TaCBFIIIb-23.1</i>		1	986	64-798	Oui	244
<i>TaCBFIIIb-23.2</i>		1	950	46-780	Oui	244
<i>TaCBFIIIb-23.3</i>		1	881	27-758	Oui	243
<i>TaCBFIIIc-10.1</i>		1	953	120-839	Oui	239
<i>TaCBFIIIc-10.2</i>		1	977	113-835	Oui	240
<i>TaCBFIIIc-13.1a</i>		1	1000	142-543	Oui	135 <sup>b</sup>
<i>TaCBFIId-16.1a</i>		1	1397	97-798	Oui	233
<i>TaCBFIId-16.2</i>		1	1562	97-798	Oui	233
<i>TaCBFIId-16.3</i>		1	976	96-797	Oui	233
<i>TaCBFIId-17.1</i>		1	1252	97-969	Oui	290
<i>TaCBFIId-24.1</i>		1	820	96-789		231
<i>TaCBFIId-24.2a</i>		1	820	87-789		234
<i>TaCBFIId-24.3</i>		1	866	50-754		234
<i>TaCBFIVa-25.1</i>		1	944	76-768	Oui	230
<i>TaCBFIVa-25.2</i>		1	1032	69-761	Oui	230
<i>TaCBFIVb-21.2</i>		1	1002	226-819	Oui	197 <sup>c</sup>
<i>TaCBFIVd-26.1</i>		1	860	89-730	Oui	213
<i>TaCBFIVd-26.2a</i>		1	946	112-750	Oui	212
<i>TaCBFIId-A27</i>	JF758493	2	48304	16818-16096		240
<i>TaCBFIId-D27</i>	JF758490	2	29456	10749-10027		240
<i>TaCBFIVa-A28</i>	JF758493	2	48304	23908-23246		220
<i>TaCBFIVa-D28</i>	JF758490	2	29456	17950-17288		220
<i>TaCBFII-5.1b</i>		3 (CS,M)	771	73-750		225
<i>TaCBFIIIc-3.2b</i>		3 (CS,M)	821	80-820		246
<i>TaCBFIIIc-13.1b</i>		1,3 (N)	803	109-516		135 <sup>b</sup>
<i>TaCBFIIIc-13.1c</i>		3 (CS,M)	834	109-333		74 <sup>d</sup>
<i>TaCBFIId-15.2b</i>		3 (CS,M)	717	1-717		238

<i>TaCBFIIIa-16.1b</i>	3 (CS,M)	762	49-750	233
<i>TaCBFIIIa-A19b</i>	3 (CS,M)	735	31-735	234
<i>TaCBFIIIa-24.2b</i>	3 (CS,M)	747	40-744	234
<i>TaCBFIVa-25.3</i>	3 (CS,M)	749	39-731	230
<i>TaCBFIVb-A20b</i>	3 (CS,M)	740	54-701	215
<i>TaCBFIVd-26.2b</i>	3 (M)	746	41-679	212
<i>TaCBFIVd-9.1b</i>	3 (CS,M)	857	47-856	269
<i>TaCBFIVd-B22b</i>	3 (CS)	893	34->893	>285 <sup>e</sup>

<sup>a</sup> Nom de gène. la procédure utilisée pour nommer de nouveaux gènes *CBF* est décrite dans le texte.

Méthode d'isolation. des codes décrivant la méthode d'isolation/identification: 1. amplification par PCR à partir de librairie d'ADNc de blé Norstar; 2. extraction de séquences de la base de données non redondante au NCBI (les séquences présentées sont issues du cultivar de blé Norstar); 3. l'amplification par PCR à partir de l'ADN génomique des cultivars de blé Norstar (N), Chinese Spring (CS) ou Manitou (M).

<sup>a</sup>. Contient un changement A → G dans le codon stop qui retarde la fin de la séquence codante pour 30 bases.

<sup>b</sup>. Contient une délétion de 32 bases dans la région codant pour le domaine AP2 qui crée un changement de cadre de lecture pour le reste de la séquence codante.

<sup>c</sup>. Contient une délétion de 15 bases avant le codon stop par rapport aux autres homeologues (Badawi et al, 2007) et le résultat est une protéine CBF plus petite.

<sup>d</sup>. Contient une délétion d'une base avant la région codant pour le domaine AP2 qui crée un changement de cadre de lecture pour le reste de la séquence codante.

<sup>e</sup>. Contient une insertion de 38 bases après le domaine AP2 qui crée un frame shift pour le reste de la séquence codante.

## 2.4 IDENTIFICATION DE POLYMORPHISMES AU NIVEAU DES GÈNES *CBF* CHEZ DEUX CULTIVARS DE BLÉ PRINTEMPS

### 2.4.1 GÉNÉRATION D'AMORCES SPÉCIFIQUES ET PCR

Le génome de blé hexaploïde contient trois génomes hérités de trois ancêtres diploïdes. Les 37 séquences de gènes *TaCBF* identifiées dans une étude précédente chez le cultivar Norstar (Badawi et al. 2007) et les 28 nouveaux gènes identifiés (cette étude) représentent 27 groupes de gènes contenant de une à trois copies homéologues dans chaque groupe. Chaque groupe de gènes a été analysé à l'aide de ClustalW (<http://www.align.genome.jp/>). Des amorces sens et antisens le plus spécifiques possible ont été choisis à l'extérieur de la région codante (en amont de l'ATG et

en aval du codon d'arrêt) (Tableau S3). Une recherche BLASTN contre tous nos *CBFs* est réalisée pour les amorces choisies afin de confirmer leur spécificité. Les amorces spécifiques ont été utilisées pour l'amplification des gènes *CBFs* à partir d'ADN génomique chez deux cultivars de printemps Chinese Spring et Manitou. L'ADN génomique de blé a été extrait et quantifié selon la méthode décrite auparavant (Limin et al. 1997). Nous avons utilisé le système d'amplification Platinum Pfx en suivant les recommandations du manufacturier (Invitrogen). Les tubes (50  $\mu$ l final) contenaient 1X tampon d'amplification Pfx, 0,3  $\mu$ M de chaque dNTP, 1mM  $MgSO_4$ , 0,3  $\mu$ M de chaque amorce spécifique, 10-20 ng d'ADN génomique, 1X enhancer et 0.5U d'ADN polymérase Platinum Pfx. Le PCR est commencé avec une incubation de 120 s à 94°C et suivi de 45 cycles de 94°C pendant 20 s, 58-60°C pendant 20 s et 68°C pendant 60 s. Les résultats de PCR ont été vérifiés par gel d'agarose et comparés avec les tailles attendues.

#### 2.4.2 CLONAGE ET ANALYSES BIOINFORMATIQUES

Les fragments obtenus par PCR ont été liés au vecteur pCR®-Blunt II-TOPO®. Les produits de ligation ont été transformés dans les bactéries compétentes *E. coli* (TOP10), qui ont été étalées sur pétri LB-agar-kanamycine (50  $\mu$ g/ml) et incubées pendant 18 heures à 37°C. Quelques colonies ont été choisies pour une mini préparation de plasmides décrite plus haut. Les plasmides contenant la séquence clonée ont été digérés par enzyme de restriction *EcoRI*. Les résultats ont été évalués par électrophorèse sur gel d'agarose. Les tailles de fragments obtenues ont été vérifiées et comparées avec les tailles attendues. Les gènes avec les tailles attendues ont été séquencés et leurs séquences sont incluses dans les figures S1 et S3.

Les résultats de séquençage ont été comparés avec les séquences des gènes *CBFs* de Norstar (Badawi et al. 2007; et notre étude) sur le site de Do-It-Yourself Sequence Comparison (<http://www.proweb.org/proweb/Tools/WU-blast.html>). Lors de ces comparaisons, nous avons évalué la présence de polymorphismes tels que des transitions, transversions ainsi que des insertions / délétions de nucléotides. Les protéines de tous les gènes qui montraient des

polymorphismes ont été traduites avec BLASTX (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) ou ORF (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>) et le ou les polymorphismes sur les chaînes protéiques identifiés. La structure secondaire de toutes les protéines CBFs avec polymorphismes a été réalisée en utilisant le programme HCA ([http://www.bioserv.rpbs.jussieu.fr/RPBS/cgi-bin/Ressource.cgi?chzn\\_lg=an&chzn\\_rsrc=HCA](http://www.bioserv.rpbs.jussieu.fr/RPBS/cgi-bin/Ressource.cgi?chzn_lg=an&chzn_rsrc=HCA)). Ces structures ont été comparées avec celles de protéines CBFs de Norstar pour déterminer si les polymorphismes affectaient la structure secondaire de protéines CBF (dans le domaine AP2, sur la signature CBF ou dans le domaine C-terminal).

## CHAPITRE III

### RÉSULTATS

Les gènes *CBF* sont des régulateurs importants de la tolérance au gel chez les plantes. Dans une étude précédente, la combinaison de séquençage d'EST de blé, le criblage de banque d'ADNc et l'amplification PCR a été utilisé pour identifier 37 gènes *CBF* exprimés chez le blé hexaploïde (Badawi et al. 2007). L'analyse moléculaire a révélé que ces gènes représentent 15 groupes de gènes avec 1 à 3 homéologues (provenant des 3 génomes A, B et D du blé hexaploïde). Des analyses phylogénétiques et bioinformatiques plus globale des *CBF* présents chez les différentes espèces de blé et autres monocotylédones ont révélé que le blé devrait posséder au moins 23-25 groupes de gènes *CBFs* différents. Comme l'information sur la famille *CBF* chez le blé était incomplète, les objectifs de notre travail étaient 1) de poursuivre l'isolation de gènes *CBF* afin d'identifier le répertoire complet chez le blé hexaploïde tolérant au gel Norstar; 2) de vérifier si deux cultivars plus sensibles au gel Chinese Spring et Manitou possèdent des polymorphismes dans la région codante des gènes *CBF* qui pourraient expliquer leur plus faible tolérance.

#### 3.1 IDENTIFICATION DE NOUVEAUX GÈNES *CBF*

Plusieurs approches ont été utilisées afin d'identifier 28 gènes additionnels différents (moins de 99% d'identité entre eux dans la région codante) chez le blé hexaploïde. L'amplification PCR à partir de bibliothèques d'ADNc (représentant diverses conditions de stress au froid, salinité et sécheresse) a permis d'isoler 23 gènes *CBF* exprimés additionnels chez le cultivar Norstar. L'analyse des séquences génomiques de blé publiées sur le site du NCBI a révélé l'existence de 4 gènes *CBF* additionnels chez le cultivar Norstar. L'amplification PCR à partir d'ADN génomique des cultivars Chinese Spring et Manitou (décrit à la prochaine section) a permis d'identifier 1 gène *CBF* additionnel chez le blé hexaploïde. Pour être conforme avec la nomenclature établie auparavant, nous avons suivi les mêmes lignes directrices (Badawi et al.

2007) avec quelques ajouts. Brièvement, une identité de 95 à 99% dans la région codante entre gènes *CBF* indiquait qu'ils étaient des copies/ homéologues (des 3 génomes A, B et D du blé hexaploïde), une identité moindre que 95% indiquait qu'ils représentaient des paralogues apparus avant la divergence des génomes ancestraux du blé, et une identité supérieure à 99% indiquait qu'ils étaient des copies/ haplotypes présentes dans la population (décrit à la prochaine section). Basé sur des analyses moléculaires et de phylogénie, un numéro de gène identique (correspondant au premier gène isolé) a été donné aux gènes *CBF* démontrant une relation orthologue entre espèces et des nouveaux numéros à des gènes identifiés dans notre étude. Dans les cas où de l'information sur la localisation génomique était disponible, une lettre (A, B, D) précède le numéro du gène. Lorsque la localisation génomique n'était pas disponible, le numéro de gène *CBF* est suivi d'une identification temporaire (.1, .2 ou .3). Parmi tous les groupes de gènes qui contiennent des copies homéologues répertoriées dans notre étude, seulement deux contiennent des membres montrant des identités inférieures à 95%. Le pourcentage d'identité exclut les bases impliquées dans des transitions et transversions tandis que les régions contenant des insertions/délétions ne sont pas incluses dans l'analyse. Ils incluent *TaCBFII-5* et *TaCBFIIIc-10* qui montrent des homologies de 91.6 et 93.9%, respectivement. Des différences marquées ont été notées entre les membres du groupe II identifiés auparavant (Badawi et al. 2007) et pourraient indiquer que ces groupes contiennent des paralogues récents et/ou qui divergent à des taux différents par rapport aux autres groupes.

Pour comprendre la relation entre les nouveaux *TaCBFs* identifiés, ceux identifiés auparavant (Badawi et al. 2007) et ceux présents chez d'autres espèces de la sous-famille *Pooideae*, nous avons compilé une série de séquences *CBF* provenant de la base de données non redondante (NR) du NCBI. Comme il est difficile d'aligner avec certitude les séquences codantes entières de membres *CBF* distants, nous avons réalisé l'analyse en deux étapes : 1) l'alignement des séquences nucléotidiques codant pour le domaine AP2 et les séquences signatures avoisinantes pour associer les membres *CBF* avec des groupes spécifiques; et 2) l'alignement séparé des séquences codantes entières de membres *CBF* des groupes apparentés *CBFIIIa*, b, c et d et des groupes apparentés *CBFIVa*, b, c et d.



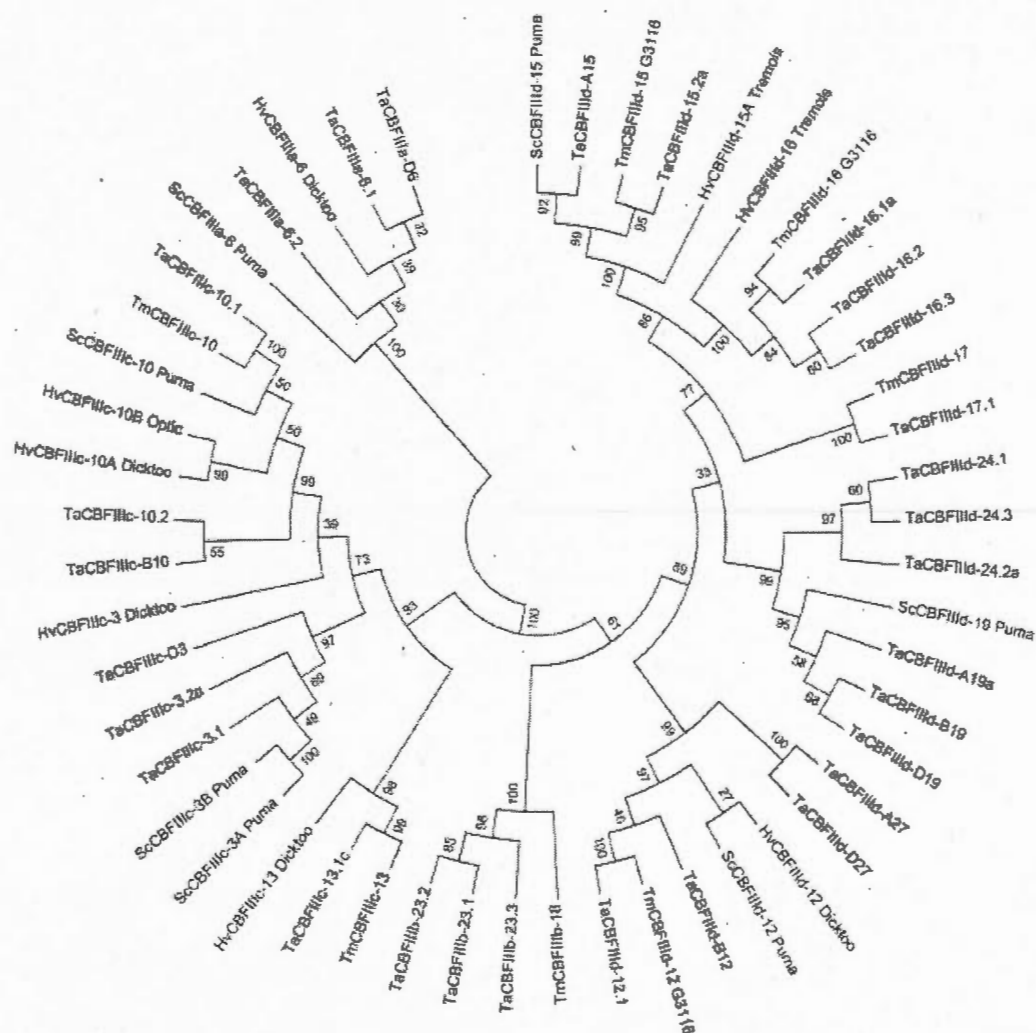
L'analyse phylogénétique des séquences nucléotidiques codant pour le domaine AP2/ régions avoisinantes révèlent la relation entre les 65 gènes *CBF* de *T. aestivum* et ceux de *T. monococcum* (10 gènes), *H. vulgare* (15 gènes) et *S. cereale* (14 gènes) (Figure 1). Cette analyse montre que le blé contiendrait 27 groupes de gènes (avec 1 à 3 homéologues chaque provenant des 3 génomes du blé hexaploïde). Nous assumons ici que le blé hexaploïde possèdera un orthologue du gène *TmCBFIIIb-18* (démontre seulement 91.3% avec le plus proche homologue *TaCBFIIIb-23.1*) et que les 5 gènes identifiés du groupe II représentent au moins deux paralogues *CBF*. Globalement les groupes I, II, IIIa, IIIb, IIIc, IVa, IVb et IVc montrent une origine unique monophylétique. Dans le cas des gènes *CBFIIIId-17* et *CBFIVd-26*, ils ne montrent pas une association (clustering) avec leurs groupes respectifs IIIId et IVd mais cette topologie est faiblement supportée (bootstrap). En ordre décroissant les groupes qui contiennent un nombre élevé de paralogues sont : groupe IIIId (7 gènes), IVd (4 gènes), IIIc et IVa (3 gènes chaque), groupes I, II, IIIb et IVb (2 gènes chaque), et groupes IIIa et IVc (1 gène chaque).



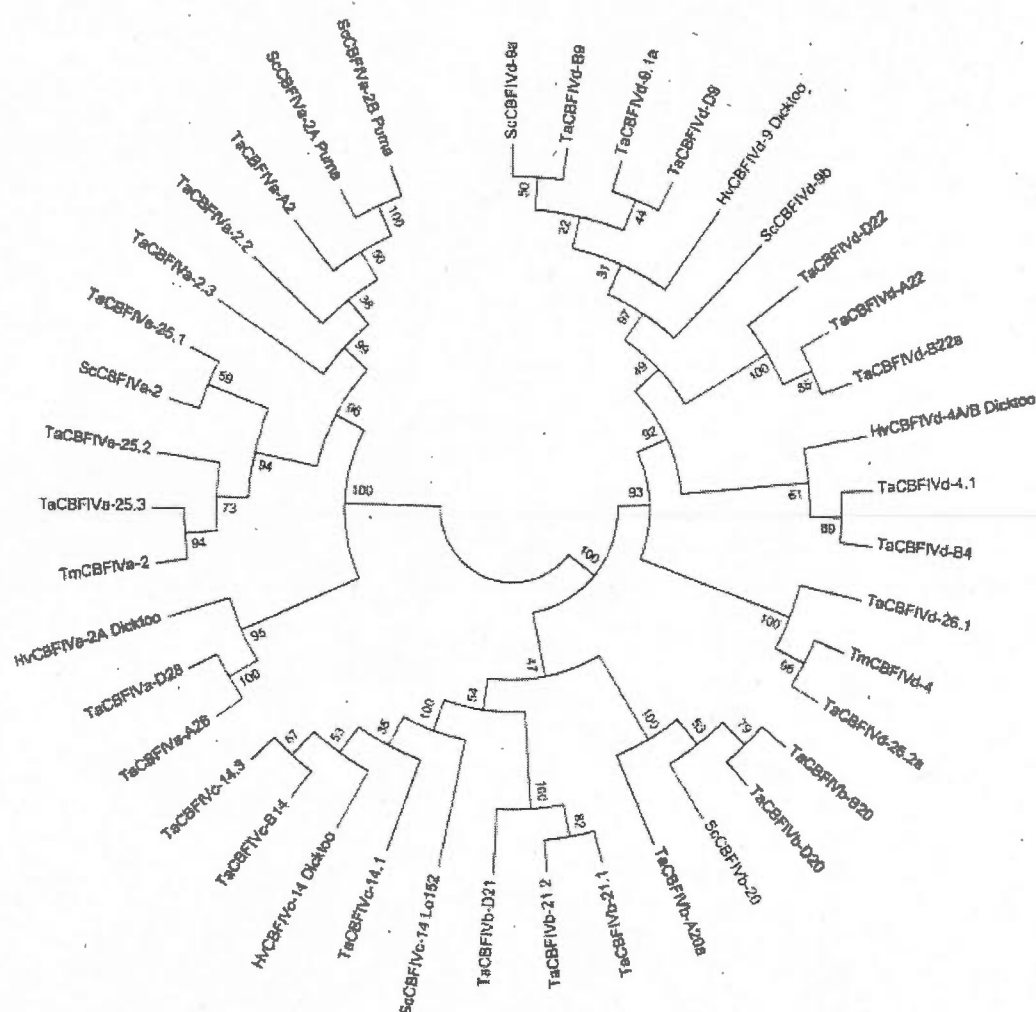
plus longs. L'analyse phylogénétique des alignements *CBF* des groupes III et IV est présentée aux Figures 2 et 3, respectivement. Ces analyses montrent que les groupes IIIa, b, c, et d (Fig 2) et IVa et d (Fig 3) sont clairement séparés. Toutefois, les gènes du groupe IVb ne sont pas entièrement séparés du groupe IVc (faible support bootstrap). Il est à noter que cette analyse a permis de montrer l'association des gènes *CBFIIIc-17* et *CBFIVd-26* avec leur groupe respectif. Il est possible de noter que les noms de gènes *CBF* sont différents entre les espèces *T. monococcum*, *H. vulgare* et *S. cereale* au niveau des groupes IVa et IVd et ceci s'explique par le choix de nom de gènes lorsque le répertoire *CBF* du groupe en question était peu représenté.

### 3.2 ANALYSE BIOINFORMATIQUE DES PROTÉINES CBF

Les gènes *TaCBF* identifiés dans notre étude codent pour des protéines ayant entre 197 et 290 acides aminés. L'analyse de leur région codante montre que certains *CBF* possèdent des différences au niveau de leur région codante qui altèrent la protéine CBF. Le gène *TaCBFII-5.5* possède une transition A→G qui détruit le codon d'arrêt et ainsi cause une extension de 10 acides aminés de la protéine vis-à-vis les autres membres du groupe *CBFII*. Dans le gène *TaCBFIVb-21.2*, une délétion de 15 nucléotides enlève les 5 derniers acides aminés de la protéine vis-à-vis les autres homéologues du gène *TaCBFIVb-21*. Dans le gène *TaCBFIIIc-13.1a*, une délétion de 32 nucléotides dans la région codant pour le domaine AP2 cause un changement de cadre de lecture et ainsi une terminaison prématurée de la protéine. Il n'est pas connu comment, ces changements affectent l'activité de ces protéines, mais pour au moins *TaCBFIIIc-13.1a*, il possède les caractéristiques d'un pseudogène.



**Figure 3.2** La relation phylogénétique entre les gènes *CBF* des groupes IIIa/b/c/d chez les Triticeae. 51 séquences nucléotidiques codant pour la région codante complète ont été alignées en utilisant ClustalW et ajusté manuellement. Un arbre a été dérivé de cet alignement (585 positions) par maximum de vraisemblance en utilisant le paramètre Tamura-3 avec bootstrap (500 répliquions). La nomenclature *CBF* est décrite dans le texte et les séquences et les numéros d'accension sont trouvés dans les figures S1 et S2. Hv *Hordeum vulgare*, Sc *Secale cereale*, Ta *Triticum aestivum*, Tm *Triticum monococcum*.



**Figure 3.3 La relation phylogénétique entre les gènes *CBF* des groupes IVa/b/c/d chez les Triticeae.** 40 séquences nucléotidiques codant pour la région codante complète ont été alignées en utilisant ClustalW et ajusté manuellement. Un arbre a été dérivé de cet alignement (657 positions) par maximum de vraisemblance en utilisant le paramètre Tamura-3 avec bootstrap (500 réplifications). La nomenclature CBF est décrite dans le texte et les séquences et les numéros d'accèsion sont trouvés dans les figures S1 et S2. Hv *Hordeum vulgare*, Sc *Secale cereale*, Ta *Triticum aestivum*, Tm *Triticum monococcum*

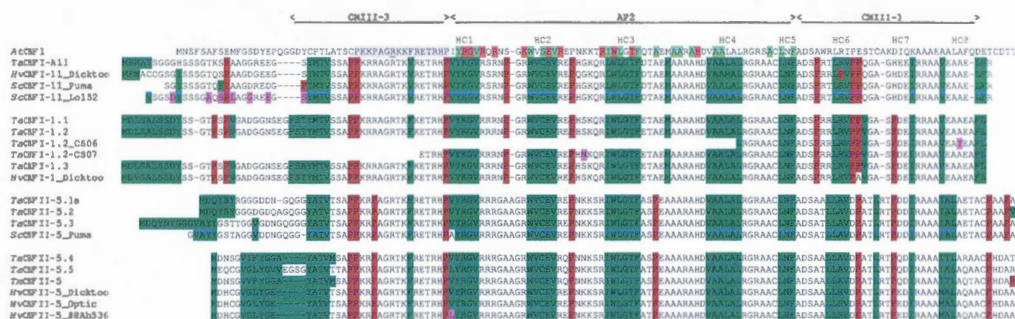
Dans l'étude précédente de notre laboratoire (Badawi et al. 2007), nous avons utilisé l'analyse des clusters hydrophobiques (HCA) afin de mettre en évidence les modifications aux faces internes de structures secondaires (hélice- $\alpha$  ou feuillet- $\beta$ ) qui pourraient être importantes



pour l'activité fonctionnelle des CBF. Le but était d'évaluer si les CBF avaient des structures communes (fonctions redondantes) ou des structures différentes (fonctions spécifiques). Cette étude a révélé que huit des dix groupes montraient des différences structurelles dans la région de liaison à l'ADN tandis que tous les dix groupes montraient des différences au niveau du domaine d'activation en C-terminal. Wang et ses collègues (2005) avaient déjà noté l'existence de clusters hydrophobiques dans le domaine d'activation C-terminal des CBF et relevé leur importance fonctionnelle. Donc dans le but d'étendre notre analyse aux nouveaux gènes identifiés dans l'étude présente et de vérifier la conservation des structures spécifiques associées aux groupes CBF chez différents espèces de *Pooideae*, nous avons compilé une série de séquences CBF provenant de la base de données non redondante (NR) du NCBI. Pour les gènes *CBFIIIb-18* et -23, *CBFIIIc-27*, -17 et -24, *CBFIVb-21*, *CBFIVd-26* et -22, aucun orthologue provenant d'autres espèces de *Pooideae* que le blé n'a été identifié à ce jour ne permettant pas ainsi d'évaluer la conservation de structure sur une plus longue période d'évolution. Deux régions ont été analysées et incluent la région entourant le domaine de liaison à l'ADN AP2 et la région entourant le domaine d'activation en C-terminal (Figures 4-9). Pour les groupes I, II, IIIa et IVc, il est possible de voir que le nombre/position/longueur des clusters hydrophobiques (HC) et le positionnement des prolines (P) est très bien conservés dans la région AP2 chez les membres analysés. Toutefois, dans la région C-terminale (Figure 5), les groupes I et II montrent des longueurs/positionnement différentes de leurs HC dans la région HC5/6 qui est spécifique aux gènes des groupes I (11 est différent de 1) et II (5.1, 5.2 et 5.3 sont différents de 5.4 et 5.5). Chez le groupe II, cette différence est une indication supplémentaire suggérant que ce groupe contient des paralogues récents. Chez le groupe IIIc au niveau du domaine AP2 (Figure 6), les protéines CBF3 et 10 montrent un HC3 plus petit et un résidu hydrophobique spécifique (I/V) entre HC3 et HC4 tandis que les protéines CBF13 montrent un HC3 élargi et n'ont pas ce résidu hydrophobique entre HC3 et HC4. De plus, CBF10 possède le plus petit HC7 du groupe IIIc. Dans leur portion C-terminale, ces protéines montrent une structure assez similaire. Dans le cas du groupe IIId, seulement de légères variations dans la longueur des HC ont été détectées en aval de la région AP2 (HC7/8, Figure 6). Toutefois, au niveau de la région C-terminale, les protéines CBF12 et 27 diffèrent des autres membres du groupe par un nombre plus élevé de HC mais de



taille réduite. Dans le cas du groupe IVa, l'analyse montre que les trois paralogues *CBF-2*, *-25* et *-28* possèdent des structures différentes au niveau de leur région AP2 qui est caractérisée par des longueurs variées de leur HC2/3 et du patron de P (Figure 8). Au niveau de leur région C-terminale (Figure 9), seul *CBF-28* semble démontrer un HC6 de longueur plus courte. Dans le cas de *CBFIVd-26*, les protéines montrent une structure qui s'apparente au groupe IVc dans la région AP2 (Figure 8). La différence majeure est que *CBF-26* possède un HC3 élargi unique. Toutefois, dans la région C-terminale (Figure 9), les protéines du groupe de gènes *CBFIVd-26* montrent une structure qui s'apparente aux autres membres du groupe IVd. La différence majeure est que *CBF-26* possède un HC3/HC4 fusionné unique. Les autres membres du groupe IVd ne montrent pas de différence majeure dans la région AP2 et C-terminale. Ces analyses suggèrent que pour refléter les différences structurales observées, plusieurs groupes pourraient être encore subdivisés en deux (groupes I, II, IIIId et IVd) et en trois (groupes IIIc et IVa). Ces résultats signifient que le blé pourrait avoir au moins 18 groupes CBF démontrant des propriétés fonctionnelles différentes.



**Figure 3.4** L'analyse comparative des noyaux hydrophobes (HC) présent dans le domaine AP2 et les signatures avoisinante des protéines CBF des Groupes I et II d'orge (*Hordeum vulgare*), blé (*Triticum spp.*) et seigle (*Secale cereale*).

Les espaces (-) ont été introduits pour maximiser l'alignement entre tous les groupes analysés.

AP2- Le domaine protéique de liaison à l'ADN.

CHIII-3 - La signature CBF en amont du domaine AP2.

CHIII-1 - La signature CBF en aval du domaine AP2.

■, Représente les noyaux d'acides aminés hydrophobes (HC) sur la chaîne protéique CBF.

■, Représente une proline.

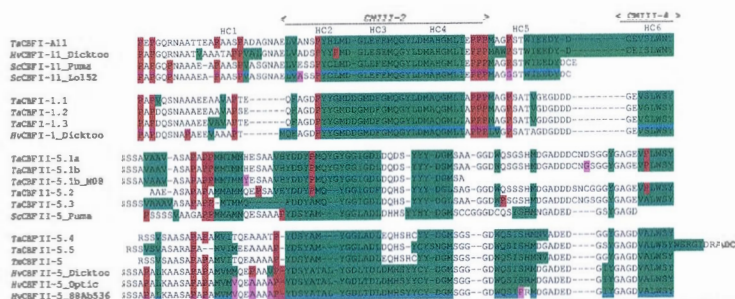
■, Représente des positions non conservées dans un homéologue chez le blé.

■, Représente les polymorphismes entre cultivars d'orge ou de blé.

■, Représente sur la séquence AtCBF1 les acides aminés des feuillets  $\beta$  qui ont une liaison avec l'ADN (Allen et al. 1998).

■, Représente sur la séquence AtCBF1 des acides aminés qui seraient impliqués dans la stabilisation hydrophobe de la structure du domaine AP2 (Allen et al. 1998).

■, Représente sur la séquence AtCBF1 des acides aminés qui seraient important pour la liaison à l'ADN (Canella et al. 2010). Des changements conservateurs sur le résidu souligné diminuent fortement la capacité de la protéine à lier l'ADN.



**Figure 3.5** L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine C-terminal des protéines CBF des Groupes I et II d'orge (*Hordeum vulgare*), blé (*Triticum spp.*) et seigle (*Secale cereale*).

Les espaces (-) ont été introduits pour maximiser l'alignement entre tous les groupes analysés.

CMIII-2 et 4 - Deux motifs du domaine d'activation C-terminal.

, Représente les noyaux d'acides aminés hydrophobes (HC) sur la chaîne protéique CBF.

, Représente une proline.

, Représente des positions non conservées dans un homéologue chez le blé.

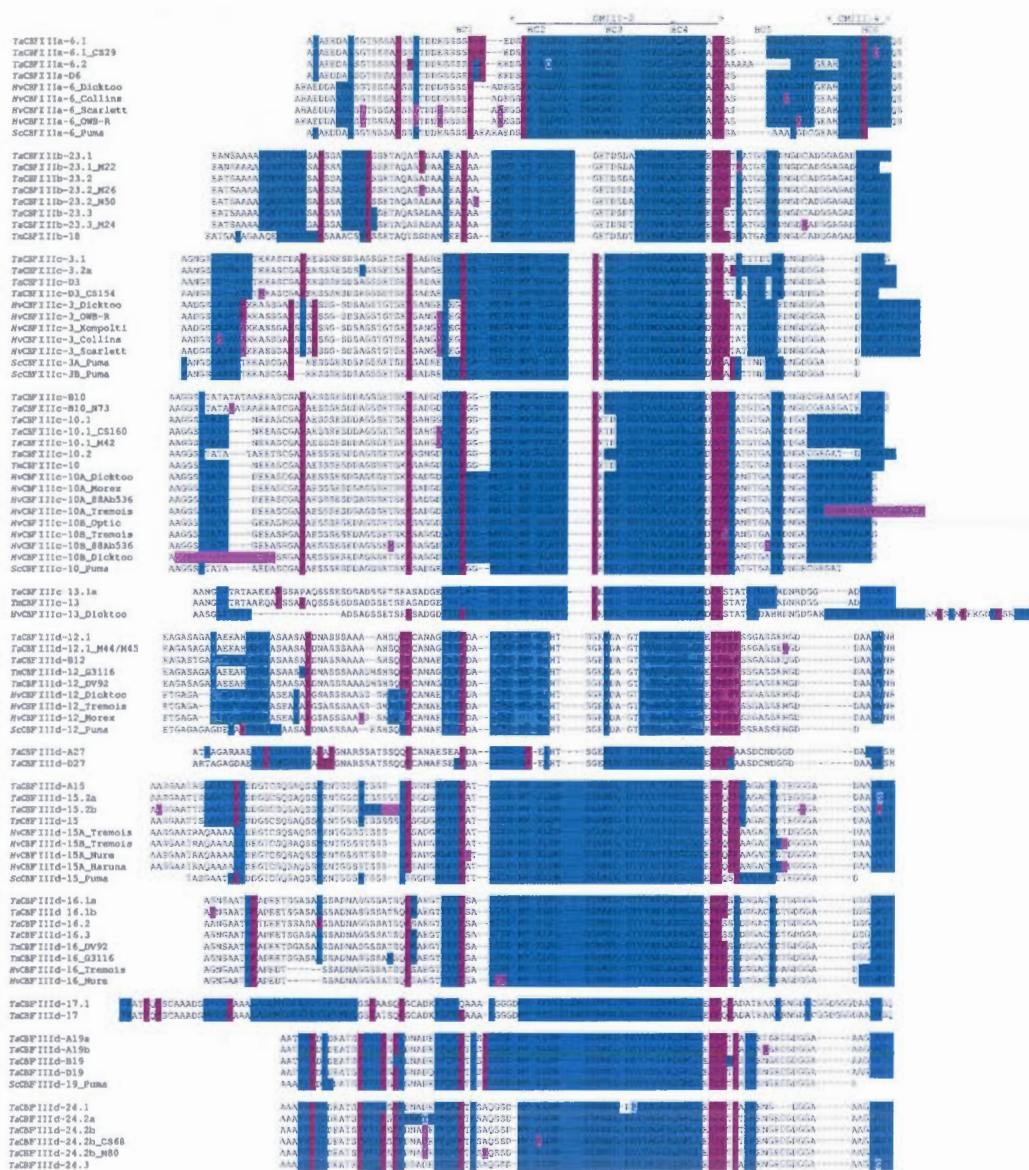
, Représente les polymorphismes entre cultivars d'orge ou de blé.



**Figure 3.6** L'analyse comparative des noyaux hydrophobes (HC) présent dans le domaine AP2 et les signatures avoisinante des protéines CBF des Groupes III d'orge (*Hordeum vulgare*), blé (*Triticum spp.*) et seigle (*Secale cereale*).

\*La description et les symboles sont décrits à la figure 4.

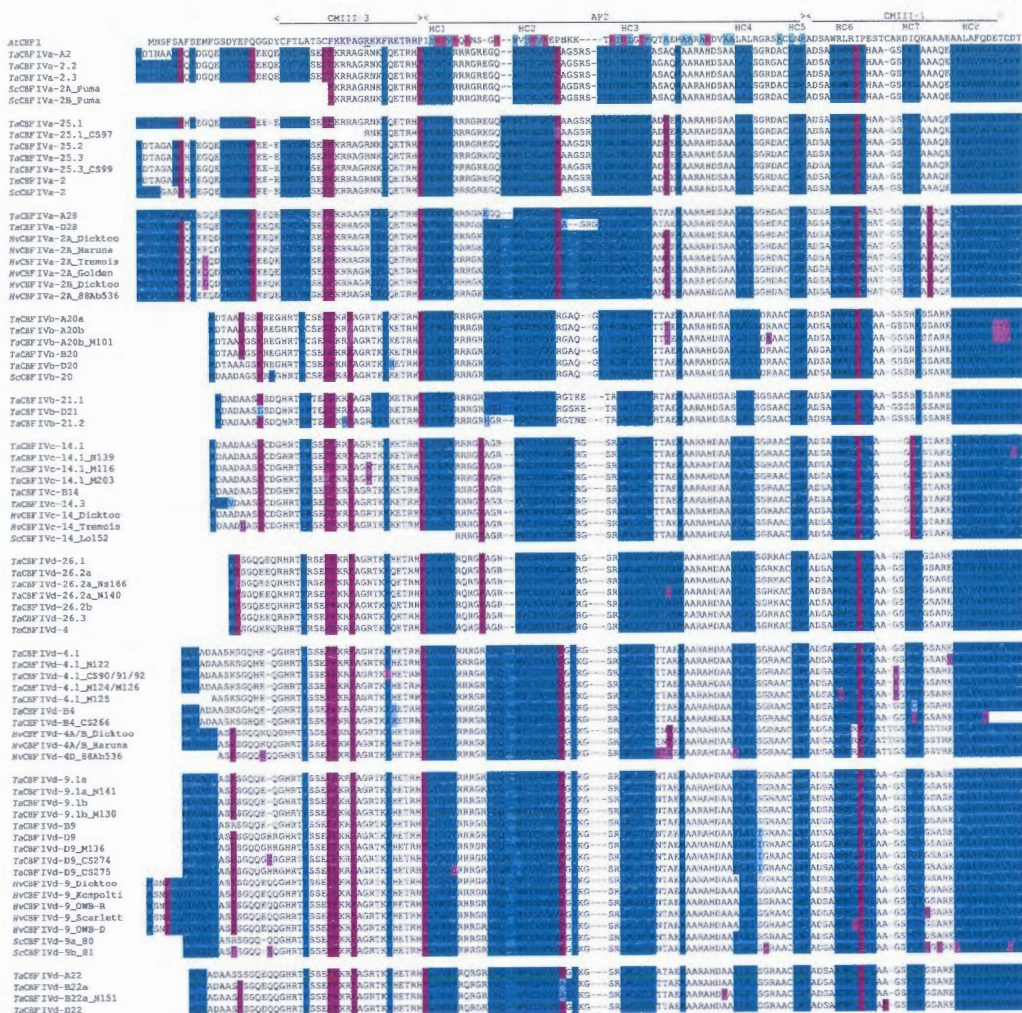




**Figure 3.7** L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine C-terminal des protéines CBF des Groupes III d'orge (*Hordeum vulgare*), blé (*Triticum spp.*) et seigle (*Secale cereale*).

\*La description et les symboles sont décrits à la figure 5.

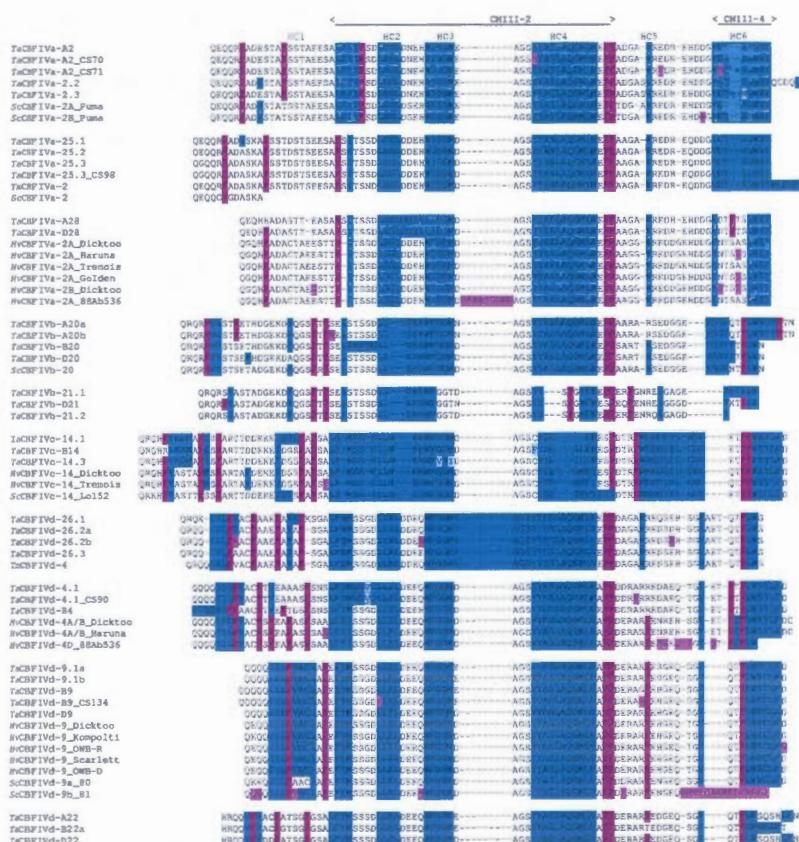




**Figure 3.8A** L'analyse comparative des noyaux hydrophobiques (HC) présent dans le domaine AP2 et les signatures avoisinante des protéines CBF des Groupes IV d'orge (*Hordeum vulgare*), blé (*Triticum spp.*) et seigle (*Secale cereale*).

\*La description et les symboles sont décrits à la figure 4.

\*La description et les symboles sont décrits à la figure 4.



\*La description et les symboles sont décrits à la figure 5.

Le blé hexaploïde possède jusqu'à 3 copies/homéologues de chaque gène *CBF* sur les génomes (A, B et D). L'analyse HCA (Figures 4-9) a été utilisée afin d'évaluer si toutes les copies des protéines CBF avaient une séquence et structure conservée (fonctions similaires et allèles optimaux) ou déviaient des autres copies ou protéines CBF, permettant ainsi d'inférer un effet sur leur activité biologique (fonctions altérées et possibilité d'allèles sub optimaux). L'accumulation de mutations (polymorphismes) peut diminuer l'activité d'une des trois copies du blé hexaploïde sans pour autant avoir un effet létale pour la plante. L'identification de ces allèles sub optimaux permettrait de répertorier les haplotypes qui pourraient être améliorés pour augmenter la dose de gènes codant pour des protéines avec une activité optimale. Pour alléger les résultats de l'analyse HCA, nous avons mis en relief seulement i) des substitutions d'acides aminés situées dans des positions hautement conservées au niveau des domaines AP2 et d'activation en C-terminal; ii) et des substitutions d'acides aminés qui affectent la longueur des HC situés dans le domaine AP2. Du côté de la région AP2, plusieurs protéines montraient des différences au niveau de la longueur de leur HC et/ou des substitutions d'acides aminés dans des positions conservées. Ces différences sont répertoriées dans le tableau 2. Du côté de la région C-terminale, plusieurs protéines montrent des substitutions dans des positions conservées (Tableau 2). Cette analyse montre que la séquence/structure de 19 copies (homéologues) chez le blé hexaploïde dévie de celle conservée dans le groupe et pourrait donc avoir un impact sur l'activité biologique de ces CBF et représenter des allèles sub optimaux.

En conclusion, l'analyse du répertoire *CBF* présent chez le cultivar Norstar, tolérant au gel, montre qu'il est complexe (au moins 64 gènes identifiés). Il contient : 1) deux pseudogènes *CBF* exprimés (*IIIc-13.1a* et *IVd-B22a*) mais non fonctionnelle à cause de la présence d'insertion/délétion dans la séquence et de codon stop interne; 2) 5 gènes *CBF* (*II-5.5*, *IIIc-3.1*, *IVa-2.2*, *IVb-A20* et *IVb-21.2*) démontrent des extensions/délétions d'acides aminés dans leur domaine C-terminal; 3) 19 gènes additionnels possèdent des changements d'acides aminés déviant des séquences/structures consensus observées.



### 3.3 IDENTIFICATION DE POLYMORPHISMES CHEZ LES CULTIVARS DE PRINTEMPS CHINESE SPRING ET MANITOU

Nous savons d'un côté que plusieurs groupes CBF (IIIc, IIId, IVa, IVb, IVc et IVd) sont associés à la tolérance au gel. Les gènes de ces groupes sont exprimés plus fortement chez les cultivars d'hiver en comparaison aux cultivars de printemps (Badawi et al. 2007). D'un autre côté, un QTL de tolérance au gel coïncide avec la localisation chromosomique des gènes *CBF* chez le blé (Miller et al. 2006; Båga et al. 2007) et chez l'orge (Skinner et al. 2006). Comme les *CBF* peuvent être à la base de cet important phénotype, une hypothèse plausible pour expliquer leur contribution pourraient être que les gènes *CBF* possèdent des différences moléculaires au niveau des régions codantes qui influencent l'activité biologique des protéines CBF exprimées. Comme l'information entourant la région codante des *CBFs* est assez exhaustive, nous avons initié l'examen de ces régions chez plusieurs cultivars. Nous avons dessiné des amorces spécifiques qui ont servi à l'amplification de gènes *CBFs* chez deux cultivars de blé de printemps plus sensible au gel, Chinese Spring et Manitou, et aussi le cultivar Norstar d'où provient à l'origine l'information de séquence.

**Tableau 3.2** Liste des homéologues CBF qui contiennent des mutations qui affectent la séquence/structure consensus CBF.

Gène	Mutation	Conséquence	Région (effet)
<i>TaCBFIIIId-24.2a</i>	c.129G>C	p.E43D	CMIII-3 (position conservée)
<i>TaCBFIIIId-24.3</i>	c.139C>T	p.P47S	CMIII-3 (position conservée)
<i>TaCBFIVb-D20</i>	c.92A>G	p.K31R	CMIII-3 (position conservée)
<i>TaCBFIVb-21.2</i>	c.64A>T	p.R22W	CMIII-3 (position conservée)
<i>TaCBFIVd-B4</i>	c.104A>G	p.H35R	CMIII-3 (position conservée)
<i>TaCBFIIIa-6.2</i>	c.317C>T	p.A106V	AP2 (HC4 étendu)
<i>TaCBFIIIc-3.1</i>	c.239C>A	p.A80D	AP2 (position conservée)
<i>TaCBFIIIc-B10</i>	c.269G>T	p.R90L	AP2 (HC5 étendu)
<i>TaCBFIIIId-D27</i>	g.242T>C	p.V81A	AP2 (HC3 raccourcit)
<i>TaCBFIIIId-A15</i>	c.238C>T	p.H80Y	AP2 (HC4 étendu)
<i>TaCBFIIIId-A19a</i>	c.221G>A	p.R74Q	AP2 (position conservée)
<i>TaCBFIIIId-24.1</i>	c.242T>C	p.V81A	AP2 (HC3 raccourcit)
<i>TaCBFIIIId-24.2a</i>	c.259G>A	p.A87T	AP2 (position hydrophobique)
<i>TaCBFIVa-A28</i>	g.176TC>AA	p.V59E	AP2 (HC2 raccourcit)
<i>TaCBFIVa-D28</i>	g.209T>C	p.V70A	AP2 (HC3 raccourcit)
<i>TaCBFIVb-21.2</i>	c.136T>C	p.Y46H	AP2 (HC2 raccourcit)
<i>TaCBFIVd-D9</i>	c.274TC>AT	p.S92I	AP2 (HC4 raccourcit)
<i>TaCBFIVd-B22a</i>	c.181C>G	p.P61A	AP2 (HC2 raccourcit)
<i>TaCBFIIIa-6.2</i>	c.392C>T	p.A131V	CMIII-1 (HC7 étendu)
<i>TaCBFIIIc-3.2a</i>	c.407T>C	p.F136S	CMIII-1 (HC8 raccourcit)
<i>TaCBFIIIId-D27</i>	g.359T>C	p.L120P	CMIII-1 (HC7 raccourcit)
<i>TaCBFIIIId-A15</i>	c.331T>C	p.S111P	CMIII-1 (HC7 raccourcit)
<i>TaCBFIVd-B4</i>	c.346G>T	p.G116C	CMIII-1 (position conservée)
<i>TaCBFIIIa-6.2</i>	c.581T>A	p.L194Q	CMIII-2 (position conservée)
<i>TaCBFIIIId-24.1</i>	c.589G>C	p.A197P	CMIII-2 (HC4 séparé)
<i>TaCBFIIIId-24.2a</i>	c.505G>T	p.D169Y	CMIII-2 (HC1 étendu)
<i>TaCBFIVc-B14</i>	c.539G>T	p.C180F	CMIII-2 (HC4 étendu)
<i>TaCBFIVc-14.3</i>	c.518G>T	p.G173V	CMIII-2 (position conservée)
<i>TaCBFIVc-14.3</i>	c.525G>C	p.M175I	CMIII-2 (position conservée)
<i>TaCBFIVd-4.1</i>	c.491G>T	p.G164V	CMIII-2 (HC2 étendu)
<i>TaCBFIIIId-15.2a</i>	c.716G>C	p.W239S	CMIII-4 (position conservée)
<i>TaCBFIIIId-24.3</i>	c.694T>G	p.W232G	CMIII-4 (position conservée)
<i>TaCBFIVa-2.3</i>	c.665C>G	p.S222W	CMIII-4 (position conservée)

Les séquences décrites dans le tableau 1 et dans (Badawi et al. 2007) ont été alignées et analysées par HCA (Figures 4-9). Seulement les substitutions survenant dans des motifs conservés (CMIII-3, AP2, CMIII-1, CMIII-2 et CMIII-4) des protéines CBF sont énumérées dans ce tableau, et elles sont regroupées en fonction de la région concernée. La nomenclature utilisée pour décrire les mutations et leur conséquence est basée sur les suggestions décrites par den Dunnen et Antonarakis (2000). Le A du codon initiateur ATG de l'homéologue présenté est noté comme nucléotide +1. Brièvement au niveau nucléotidique, le G présent dans les autres homéologues est remplacé par C (position 129) dans *TaCBFIIIId-24.2a* tandis qu'au niveau protéique, le E dans les autres homéologues est remplacé par D (position 43) dans *TaCBFIIIId-24.2a*.



À ce jour, notre première série d'oligos a permis de ré-amplifier spécifiquement des variantes de 50 des 65 gènes *CBF* identifiés chez Norstar. Ces gènes montrent soit 100% d'homologie avec la séquence initiale de Norstar ou soit, ils représentent un proche variant (>99%) du gène (Tableau 3). Globalement, l'analyse a permis de reconfirmer la séquence initiale Norstar pour 41 des 50 gènes ré-amplifiés. Les cultivars de printemps partagent en grande partie les mêmes variantes/haplotypes avec Norstar car 31 des 41 gènes reconfirmés ont été trouvés chez au moins un génotype de printemps. Le fait que nous n'avons pas confirmé toutes les séquences initiales est dû à plusieurs raisons, dont une priorisation pour des combinaisons (amorces-ADN blé de printemps), le nombre limité de séquences analysées (239 séquences), et/ou l'existence d'erreurs dans la séquence initiale de Norstar. D'autre part, cette analyse a permis de révéler des haplotypes (>99% d'homologie) de 10 gènes *CBF* retrouvés spécifiquement chez les cultivars de printemps (9 de ces gènes ont été confirmés dans les deux cultivars). Ces gènes ont été nommés en introduisant une lettre minuscule additionnelle après le numéro du gène et incluent : *TaCBFII-5.1b*, *TaCBFIIIc-3.2b*, *TaCBFIIIc-13.1c*, *TaCBFIIId-15.2b*, *TaCBFIIId-16.1b*, *TaCBFIIId-A19b*, *TaCBFIIId-24.2b*, *TaCBFIVb-A20b*, *TaCBFIVd-9.1b* et *TaCBFIVd-26.2b*. De plus, cette analyse a permis d'identifier un variant/copie additionnelle chez Norstar *TaCBFIIIc-13.1b* (>99% d'homologie) et un nouvel homologue identifié chez Chinese Spring et Manitou *TaCBFIVa-25.3* (98.1% d'homologie versus *TaCBFIVa-25.2*). Le reste des variantes de séquence présentées au Tableau 3 représente potentiellement des haplotypes spécifiques à un cultivar et/ou des évidences de copies additionnelles du gène chez un cultivar. On peut dénombrer 30, 21, 13 copies paralogues/haplotypes spécifiques chez Chinese Spring, Manitou et Norstar, respectivement. Comme ces informations n'ont pas été vérifiées chez d'autres cultivars, ces séquences n'ont pas été déposées dans la base de données du NCBI non redondante, mais sont rajoutées à l'analyse pour déceler si des variations dans leur séquence affectent potentiellement leur structure et/ou activité biologique.

Tableau 3.3 Liste de tous les gènes/paralogues/haplotypes de CBF retrouvés chez le blé hexaploïde.

Gène CBF	100% homologue avec CDS de gène Norstar initial	Norstar (spécifique)	Chinese Spring (spécifique)	Manitou (spécifique)
<i>TaCBFI-1.1</i>	CS (541/654) Man (585/654)			
<i>TaCBFI-1.2</i>	Man (654/654)		<u>S06 (353/654)</u> <b>G&gt;A 412, A&gt;T 612</b> <b>A138T</b>  <u>S07 (508/654)</u> <b>G&gt;A 227, A&gt;T 612</b> <b>S76N</b>	
<i>TaCBFI-1.3</i>	Nor (654/654) CS (488/654) Man (484/654)			
<i>TaCBFII-5.1b</i>	CS (678/678) Man (577/678)  <u>S11+M09 versus 5.1a</u> <b>A&gt;G 632b</b> <b>D211G</b>			<u>M08 (576/678)</u> <b>C&gt;T469</b> <b>H157Y</b>
<i>TaCBFII-5.2</i>	Man (660/660) CS (593/660) Nor (584/660)			
<i>TaCBFII-5.3</i>	CS (687/687) Nor (522/687)			<u>M14 (436/687)</u> <b>C&gt;T432b</b>
<i>TaCBFII-5.4</i>	CS (633/633) Nor (419/633)			
<i>TaCBFII-5.5</i>	Man (409/675) CS (524/675) Nor (675/675)			
<i>TaCBFIIIa-6.1</i>	Man (711/711)	<u>N33 (711/711)</u>	<u>S29 (638/711)</u>	<u>M21 (711/711)</u>

		<b>A&gt;G 86</b> <b>E29G</b>	<b>G&gt;T 696</b> <b>W232C</b>	<b>G&gt;A 376</b> <b>A126T</b>
<i>TaCBFIIIa-6.2</i>		<u>N38 (711/711)</u> <b>T&gt;A 546</b>		
<i>TaCBFIIIa-D6</i>	Nor (717/717) Man (602/717)			
<i>TaCBFIIIb-23.1</i>	Man (703/735)		<u>S197 (581/735)</u> <b>A&gt;T 252</b>	<u>M22 (726/735)</u> <b>T&gt;C 30, G&gt;A 655</b> <b>V219M</b>  <u>M27 (572/735)</u> <b>ΔG 61</b>
<i>TaCBFIIIb-23.2</i>	Man (666/735) CS (735/735) Nor (694/735)	<u>N50 (607/735)</u> <b>G&gt;A 544</b> <b>A182T</b>	<u>S195 (683/735)</u> <b>G&gt;A 327</b>	<u>M26 (518/735)</u> <b>G&gt;C 517</b> <b>A173P</b>
<i>TaCBFIIIb-23.3</i>				<u>M24 (524/732)</u> <b>T&gt;C 685</b> <b>C229R</b>
<i>TaCBFIIIc-3.1</i>	Man (591/732) CS (656/732) Nor (687/732)			<u>M159 (600/732)</u> <b>G&gt;A 132</b>
<i>TaCBFIIIc-3.2a</i>	Nor (683/741)			
<i>TaCBFIIIc-3.2b</i>	CS (694/741) Man  <u>S217 et M165</u> Versus <i>TaCBFIIIc-3.2a</i> <b>C&gt;A 222, G&gt;C 263,</b> <b>G&gt;A 367, T&gt;C 407</b>  <b>Y74S,R88G, P136S</b>			
<i>TaCBFIIIc-D3</i>	Man (738/738) Nor (583/738)		<u>S153 (738/738)</u> <b>C&gt;T 237</b>	

			<u>S154 (689/738)</u> C>T 67, G>A 304, G>A 454 H23Y, A102T, E152K	
<i>TaCBFIIIc-B10</i>	CS (723/723) Man (723/723)	<u>N73 (649/723)</u> C>T 413, T>C 465 T138I		
<i>TaCBFIIIc-10.1</i>	Nor (663/723)		<u>S160 (676/720)</u> G>A 353, C>T 471, G>C 518 R118H, G173A	<u>M42 (657/723)</u> G>A 131, C>T 471, G>C 518 R44H, G173A
<i>TaCBFIIIc-10.2</i>	Man (650/723)			
<i>TaCBFIIIc-13.1b</i>	Nor (704/733)  N163 versus TaCBFIIIc-13.1a) C>G 202			
<i>TaCBFIIIc-13.1c</i>	CS (685/733) Man (583/733)  S221,S222,S224 M175,M176,M177, M178 versus TaCBFIIIc-13.1a C>A 307		<u>S223 (632/733)</u> C>T 358	
<i>TaCBFIIIc-12.1</i>	Nor (397/738)		<u>S227 (522/738)</u> G>A 72	<u>M44 (738/738)</u> G>C 612, G>A 707 R236H
<i>TaCBFIIIc-B12</i>	Man (520/738) CS (559/738)		<u>S24 (489/738)</u> C>T 576	

	Nor (588/738)			
<i>TaCBFIIIId-A15</i>	Man (720/720) CS (692/720) Nor (720/720)		<u>S231 (675/720)</u> <b>G&gt;A 229</b> <b>A77T</b>	
<i>TaCBFIIIId-15.2b</i>	Man (673/726) CS (717/717)  <u>S41/M56</u> versus 15.2a <b>C&gt;T 84, T&gt;C 243,</b> <b>C&gt;T 293, G&gt;A 397,</b> <b>ΔCGTCGTCGT</b> <b>503, A&gt;G 685, C&gt;G</b> <b>707</b>  <b>A&gt;V28, M&gt;T 81,</b> <b>P&gt;S 97, ΔRRR 168,</b> <b>R&gt;G235</b>			
<i>TaCBFIIIId-16.1a</i>	Nor (702/702)			
<i>TaCBFIIIId-16.1b</i>	Man (702/702) CS (659/702)  <u>S44/M58</u> Versus <i>TaCBFIIIId-16.1a</i> <b>T&gt;C 400, T&gt;C 507</b> <b>S134P</b>			
<i>TaCBFIIIId-16.2</i>	Man (657/702) CS (380/702) Nor (702/702)			
<i>TaCBFIIIId-A19a</i>		<u>N97 (658/705)</u> <b>G&gt;A 221</b> <b>R74Q</b>		
<i>TaCBFIIIId-A19b</i>	Man (694/705)			



	CS (705/705)  <u>S55/M65</u> Versus <i>TaCBFIIIId-A19a</i> <b>C&gt;T 86, G&gt;T 265,</b> <b>C&gt;A 656</b> <b>S29L, A89S, T219N</b>			
<i>TaCBFIIIId-B19</i>	Man (546/706) Nor (706/706) CS (684/706)			
<i>TaCBFIIIId-D19</i>	Man (539/705) Nor (705/705) CS (672/705)			
<i>TaCBFIIIId-24.2a</i>	Nor (554/705)			
<i>TaCBFIIIId-24.2b</i>	CS (705/705) Man (705/705)  <u>S64/M75</u> Versus <i>TaCBFIIIId-24.2a</i> <b>T&gt;C 30, C&gt;G 129,</b> <b>T&gt;C 381, T&gt;G 505,</b> <b>G&gt;A 654</b> <b>V&gt;A10, T&gt;R43,</b> <b>L&gt;P127</b>		<u>S68 (682/705)</u> <b>G&gt;A 556</b> <b>E186K</b>	<u>M78 (705/705)</u> <b>G&gt;A 87</b>  <u>M80 (701/705)</u> <b>C&gt;A 536</b>
<i>TaCBFIVa-A2</i>	CS (678/678) Nor (678/678)		<u>S70 (678/678)</u> <b>T&gt;C 568</b> <b>Y190H</b>  <u>S71 (678/678)</u> <b>A&gt;T 629, G&gt;A 655</b> <b>E210V, D219N</b>	<u>M81 (678/678)</u> <b>A&gt;G 213</b>
<i>TaCBFIVa-2.2</i>	Man (591/771)			

	CS (732/771) Nor (771/771)			
<i>TaCBFIVa-25.1</i>	Man (633/693) CS (693/693) Nor (602/693)		<u>S97 (584/693)</u> <b>C&gt;T 209</b> <b>P70L</b>	
<i>TaCBFIVa-25.2</i>	Nor (693/693)			
<i>TaCBFIVa-25.3</i>	Man (693/693) CS (693/693)			
<i>TaCBFIVb-A20b</i>	CS (648/648) Man (603/648)  <u>N215/M100</u> versus <i>TaCBFIVb-A20a</i> <b>C&gt;T 71, A&gt;G 93,</b> <b>G&gt;T 213, G&gt;A 214,</b> <b>C&gt;T 273, Δ CGC</b> <b>TTA 377, A&gt;C 466</b> <b>R&gt;W 24, K&gt;R31,</b> <b>R&gt;I71, P&gt;L91, ΔPL</b> <b>126</b>			<u>M101 (582/648)</u> <b>G&gt;A 266</b> <b>R89H</b>
<i>TaCBFIVb-B20</i>	Man (586/639)			
<i>TaCBFIVb-21.1</i>	Man (609/609) CS (609/609) Nor (609/609)			
<i>TaCBFIVb-D21</i>	Man (609/609) CS (589/609) Nor (609/609)			
<i>TaCBFIVb-21.2</i>	Nor (609/609)			
<i>TaCBFIVc-14.1</i>		<u>N139 (639/639)</u> <b>G&gt;A 376</b> <b>A126T</b>	<u>S107 (639/639)</u> <b>G&gt;A 576</b>	<u>M116 (639/639)</u> <b>C&gt;T 79, G&gt;A 576</b> <b>R27W</b>  <u>M203 (639/639)</u>

				C>T 59, C>T 79, G>A 576 P20L, R27W  <u>M204 (639/639)</u> C>T 79, C>T 567, G>A 576 R27W
<i>TaCBFIVc-B14</i>	Man (645/645) CS (606/645) Nor (645/645)			
<i>TaCBFIVc-14.3</i>	Man (597/645) CS (531/645)			
<i>TaCBFIVd-4.1</i>	Nor (669/669)	<u>N122 (669/669)</u> A>T 365 E122V	<u>S90 (669/669)</u> C>G 102, G>A 338, C>TG>A 338 599 F34L, G113E  <u>S91 (630/669)</u> C>G 102, G>A 338 F34L, G113E	<u>M124 (662/669)</u> G113E  <u>M125 (654/669)</u> C>T 313, G>A 338 R105W, G113E
<i>TaCBFIVd-B4</i>	Nor (670/670) Man (670/670)		<u>S266 (670/670)</u> T>G 346, 381^382insA, C>T 538 C116G, 127^128 ins S (changement de cadre de lecture)  <u>S267 (670/670)</u> T>G 346, 381^382insA, C>T538, C>T 650 C116G, 127^128 ins S (changement de cadre de lecture)	

<i>TaCBFIVd-9.1a</i>		<u>N141 (700/810)</u> A>G 494 Q165R		
<i>TaCBFIVd-9.1b</i>	CS (810/810) Man (727/810)  S110/M131 versus <i>TaCBFIVd-9.1a</i> C>A 309, G>C 393, A>T 404, C>A 538, G>T 633, G>C 747		<u>S128 (722/810)</u> A>G 789  <u>S131 (672/810)</u> C>T 249	<u>M130 (636/810)</u> T>C 404 Q135P
<i>TaCBFIVd-26.2a</i>	Nor (639/639)	<u>S165 (639/639)</u> G>T 624  <u>S166 (639/639)</u> C>A 5 P2Q  <u>N140 (639/639)</u> C>A 206 A69D		
<i>TaCBFIVd-26.2b</i>	Man (639/639)  <u>M207</u> versus <i>TaCBFIVd-26.2a</i> A>G 51, A>G 240, C>T 303, T>C 407, C>G 496, C>T 516, C>G 597 H>R17, Y>C80, P>L101, S>P135, S>R165, S>L172, T>S199			
<i>TaCBFIVd-B9</i>	Man (776/810)		<u>S132 (748/810)</u>	

	Nor (730/810)		G>A 748  S134 (740/810) T>C 644, G>A 748 L215P, G250P	
<i>TaCBFIVd-D9</i>	Nor (661/810)		S135 (793/810) C>T 657  S274 (638/810) C>T 43 H15Y  S275 (675/810) T>A 137 V46E	M136 (674/810) C>T 496 Codon stop
<i>TaCBFIVd-A22</i>	Man (724/828)			
<i>TaCBFIVd-B22</i>	Man (681/860)	N151 (782/860)	S148 (677/860) C>T 254, G>A 253, A>G 706, A>G 787	

Dans les haplotypes (b) et (c), les mutations et leurs conséquences sont rapportés en comparaison avec l'haplotype d'origine (a) de Norstar. Pour les haplotypes spécifiques détectés chez N, CS et M, les mutations et leurs conséquences sont rapportés en comparaison avec l'haplotype décrit dans la première colonne.

La nomenclature des mutations est basée sur les recommandations de Dunnen et Antonarakis, (2000).

\*T>A 144 (Substitution de T par A dans la 144<sup>e</sup> position nucléotidique)

\*H40E (Substitution de l'acide aminé H par l'acide aminé E dans la 40<sup>e</sup> position de la protéine)

\*ΔG 120 (Déletion d'un nucléotide ou acide aminé G sur la 120<sup>e</sup> position de la chaîne nucléotidique/protéique, respectivement)

\*130^131 ins. A (Insertion d'un nucléotide ou acide aminé A entre la position 130 et 131 de chaîne nucléotidique/protéique, respectivement)

En analysant le Tableau 3, il est possible de dénombrer que 110 changements d'une base (SNP) ont été détectés. Ces changements incluent 75 transitions (A↔G, T↔C) et 35 transversions (A/G↔C/T). Donc selon nos calculs, environ 70% des changements sont des changements de transition. Pratiquement dans toutes les séquences d'ADN/génome étudié, les changements de transition se produisent à des fréquences supérieures aux changements de



transversion (Brown et al. 1982; Gojobori et al. 1982; Curtis et Clegg, 1984; Wakeley, 1994, 1996).

Sur les 110 polymorphismes détectés sur les gènes *CBFs*, 43 représentent des changements silencieux, ne causant aucun changement sur les protéines, 66 représentent des mutations faux-sens causant des substitutions d'acides aminés, et une représente une mutation non-sens qui change un codon d'acide aminé par un codon stop (Tableau 4). Les mutations faux-sens comprennent des substitutions d'acides aminés hydrophobes ou hydrophiles par un autre acide aminé de la même catégorie ou de catégorie différente. On peut prévoir que des substitutions par des acides aminés qui se ressemblent au niveau biochimique auront une incidence moindre sur la fonction des protéines CBF. De plus, plusieurs gènes *CBF* contenant des insertions/délétions de bases ont été identifiés. Comme ces mutations changent le cadre de lecture, elles rendent les protéines non fonctionnelles.

Sur les 66 mutations faux-sens, 15 mutations ont été identifiées sur la chaîne protéique dans le domaine AP2 et 10 mutations dans le domaine C-terminal. En tenant compte de la longueur de la région analysée, le tableau 4 montre que le nombre de mutations trouvé sur les signatures CBF ou entre les domaines AP2 et C-terminal est plus grand que celui présent dans les domaines conservés AP2 et C-terminal (Figures 4-9).

**Tableau 3.4** Nombre de polymorphismes (SNP) causant des mutations faux- et non-sens sur les différentes parties de la protéine CBF.

Nombre de mutations sur la signature CBF CMIII-3 ( $\approx$ 25 acides aminés)	11
Nombre de mutations sur le domaine AP2 ( $\approx$ 60acides aminés)	15
Nombre de mutations sur la signature CBF CMIII-1 ( $\approx$ 30 acides aminés)	12
Nombre de mutations entre les domaines AP2 et C-terminal ( $\approx$ 25-50 acides aminés)	19
Nombre de mutations sur le domaine C-terminal ( $\approx$ 70 acides aminés)	10
Nombre total de SNP	66

Les calculs sont basés sur les figures de l'analyse HCA (Figures 4-9)

Lors des analyses HCA des protéines de deux cultivars de printemps, Chinese Spring et Manitou, nous avons dans un premier temps étudié les orthologues de gènes Norstar déviants des séquences consensus. Nous avons trouvé que les deux pseudogènes (*IIIc-13.1* et *IVd-B22*) qui existent chez Norstar, sont présents chez Manitou et Chinese Spring. Toutefois, le gène

*TaCBDIIIc-13.1c* est différent et ressemble beaucoup à son orthologue chez le blé diploïde, il contient une insertion de 11 acides aminés dans le domaine AP2, mais subit en plus une délétion d'une base (19delT en amont de CMIII-3) qui change le cadre de lecture. Parmi les cinq gènes avec des délétions/insertions dans le domaine C-terminal chez Norstar, l'information de séquence incomplète pour les gènes *TaCBFII-5.5* et *TaCBFIVc-21.2* ne permet pas de tirer des conclusions tandis que les gènes *TaCBFIIc-3.1*, *TaCBFIVa-2.2* et *TaCBFIVb-A20b* chez les cultivars de printemps, ont des régions C-terminales de longueur similaire à leur orthologue chez Norstar. Parmi les 19 protéines Norstar possédant des changements d'acides aminés dans le domaine AP2 ou C-terminal, sept gènes (*TaCBFIIa-6.2*, *TaCBFIIId-D27*, *TaCBFIIId-24.1*, *TaCBFIIId-24.3*, *TaCBFIVa-A28*, *TaCBFIVa-D28*, *TaCBFIVa-2.3*) n'ont pas encore été ré-amplifiés chez les cultivars de printemps. D'autre part, chez Chinese Spring et Manitou, huit protéines montrent les mêmes changements que les protéines *TaCBFIIc-B10*, *TaCBFIIId-A15*, *TaCBFIVc-B14* et *TaCBFIVc-14.3*, *TaCBFIVd-D9*, *TaCBFIVd-B22a*, *TaCBFIVd-4.1* et *TaCBFIVd-B4* de Norstar. Sur les quatre protéines qui diffèrent de celles de Norstar à ces positions, seulement une protéine représente une version améliorée, car elle se rapproche de la séquence consensus tandis que trois sont des versions équivalentes. Par exemple, les substitutions problématiques dans les domaines AP2 et C-terminal de *TaCBFIIId-24.2a* sont rétablies sur *TaCBFIIId-24.2b* (D43E, Y169D). La protéine *TaCBFIIc-3.2b* montre une substitution (S136F) sur le HC8 qui rétablit le consensus à cette position, mais le changement (A88G) entre les HC3 et HC4 sur l'hélice  $\alpha$  du domaine AP2 l'éloigne du consensus. La position occupée par l'alanine est importante pour la stabilisation hydrophobique de la structure du domaine AP2 (Allen et al. 1998). Malgré le changement (Q74R) chez *TaCBFIIId-A19b* qui a rétabli le consensus à cette position, le changement de l'alanine sur l'hélice  $\alpha$  (A89S) l'éloigne du consensus. Chez *TaCBFIIId-15.2b*, la substitution (S236W) sur le domaine HC6 du domaine C-terminal rétablit le consensus à cette position, mais à cause du remplacement d'acide aminé hydrophobe (A98V), les deux HC (HC5 et HC6) de la région AP2 sont devenus continuels. Pour les autres haplotypes présentés dans le Tableau 3, quatre allèles *TaCBFII-5.1b*, *TaCBFIIId-16.1b*, *TaCBFIVd-26.2b* et *TaCBFIVd-9.1b* codent pour des protéines ne présentant pas de déviations majeures par rapport au consensus. La protéine *TaCBFIVb-A20b* contient trois

changements majeurs, sur la signature CBF, CMIII-3 (P24L), dans le domaine AP2 (A72T), et sur le CMIII-1 (AV122 del) par rapport à son homéologue chez Norstar, qui ont comme effet de l'éloigner du consensus.

Donc il est clair que les gènes *CBF* chez les cultivars de printemps sont complexes. Ils contiennent les deux pseudogènes retrouvés chez Norstar et la majorité des protéines analysées présentent les mêmes changements que Norstar vis-à-vis la séquence consensus. Une des protéines analysées TaCBFIIIId-24.2b est plus similaire à la séquence consensus que la variante chez Norstar. Trois autres protéines TaCBFIIIId-3.2a, TaCBFIIIId-A19a et TaCBFIIIId-15.2a, présentent des changements qui s'annulent, certaines modifications l'approchent du consensus et d'autres l'éloignent. La dernière protéine analysée, TaCBFIVb-A20b, possède des changements qui l'éloignent du consensus.

**Tableau 3.5** Liste des mutations survenant dans des portions conservées de gènes *CBF* groupées selon la région impliquée.

Gène	Mutation (ADN)	Conséquence (Protéine)	Région
<i>TaCBFIIIb-23.1</i>	M27 (ΔG 61)	63delG	CMIII-3
<i>TaCBFIVc-14.1</i>	M116 (C>T 79) M204 (C>T 79) M203 (C>T 59, C>T 79)	R27W R27W P20L, R27W	CMIII-3 CMIII-3 CMIII-3
<i>TaCBFIIIc-10.1</i>	M42 (G>A 131)	R44H	AP2
<i>TaCBFIIId-A15</i>	S231 (G>A 229)	A77T	AP2
<i>TaCBFIVa-25.1</i>	S97 (C>T 209)	P70L	AP2
<i>TaCBFIVd-26.2a</i>	N140 (C>A 206)	A69D	AP2
<i>TaCBFIVd-D9</i>	S275 (T>A 137)	V46E	AP2
<i>TaCBFIIId-6.1</i>	M21 (G>A 376)	A126T	CMIII-1
<i>TaCBFIIIc-10.1</i>	S160 (G>A 353)	R118H	CMIII-1
<i>TaCBFIVd-4.1</i>	N122 (A>T 365) M125 (C>T 313)	E122V R105W	CMIII-1 CMIII-1
<i>TaCBFIVd-D9</i>	M136 (C>T 496)	Q166X	en aval de CMIII-1
<i>TaCBFIIId-24.2b</i>	S68 (G>A 556)	E186K	CMIII-2
<i>TaCBFIVa-A2</i>	S70 (T>C 568)	Y190H	CMIII-2
<i>TaCBFIVd-B9</i>	S134 (T>C 644)	L215P	CMIII-2

\* La liste complète des gènes et des mutations peut être trouvée dans le tableau 3. La nomenclature utilisée pour décrire les mutations et les conséquences est basée sur les suggestions décrites par den Dunnen et Antonarakis (2000). Le A du codon initiateur ATG (méthionine) de la séquence Norstar est noté nucléotide +1. Dans la deuxième colonne, la lettre indique les cultivars de blé: N, Norstar; CS, Chinese Spring; et M, Manitou, et le nombre indique le nom du clone. Les motifs conservés dans les protéines CBF (CMIII-3, AP2, CMIII-1, CMIII-2 et CMIII-4) sont indiqués dans les figures 4, 5, 6, 7, 8 et 9.

Parmi les haplotypes spécifiques retrouvés chez les trois cultivars, Manitou contient 16 haplotypes nouveaux et Chinese Spring en a 17. Le nombre de mutations non-sens/faux-sens est 17 et 20 chez chacun de ces cultivars, respectivement. Nous avons présenté dans le tableau 5 uniquement les changements se produisant dans des régions conservées, car on peut prévoir que ces changements auront plus d'effet sur la structure de la protéine et son activité biologique. Par



exemple, nous savons que les 16 acides aminés sur la signature CBF, en amont du domaine AP2 sont hautement conservés et qu'ils ont probablement une fonction spécifique inconnue qui permet la liaison de la protéine à l'ADN (Canella et al. 2010). De plus, plusieurs acides aminés critiques ont été identifiés pour la liaison à l'ADN dans le domaine AP2 (Allen et al. 1998). Parmi les trois haplotypes du gène *TaCBFIIIb-23.1* détectés chez Manitou (Tableau 3), M25 (100% homologue à Norstar), M22 et M27, seul M27 a une délétion d'un nucléotide sur le CMIII-3 en amont du domaine AP2 (Figure 6). Cette délétion cause un changement dans le cadre de lecture après la position 22. Pour le gène *TaCBFIIIc-10.1*, l'haplotype chez Manitou (M42) a une substitution (R44H) dans le domaine AP2, sur un résidu très conservé tandis que l'haplotype chez Chinese Spring (S160) montre une substitution (R118H) dans une position différente, mais aussi sur un résidu très conservé. Donc les deux cultivars de printemps ont des haplotypes du gène *TaCBFIIIc-10.1* qui s'éloignent du consensus. Pour le gène *TaCBFIVd-B4*, Chinese Spring possède deux haplotypes, S266 et S267, qui montrent une insertion d'une base dans le CMIII-1 et ainsi un changement dans le cadre de lecture qui rend leur protéine non fonctionnelle. Manitou et Norstar n'ont pas cette insertion. Pour le gène *TaCBFIVd-D9*, l'haplotype M136 de Manitou possède une substitution d'une base qui provoque l'apparition d'un codon stop entre le domaine AP2 et C-terminal. Norstar et Chinese Spring n'ont pas cette mutation. Toutefois, un des trois haplotypes détectés de Chinese Spring (S275) possède une substitution V46E qui arrive dans une position importante pour la structure/fonction du domaine AP2. Cette substitution rend aussi le HC1 du domaine AP2 plus petit. Le gène *TaCBFIVc-14.1* a deux haplotypes chez Manitou, M116 et M203. Tous les deux montrent une substitution (R>W) dans le domaine CMIII-3 dans une position très conservée et très importante pour l'interaction de la protéine CBF et le promoteur. Le gène *TaCBFIVd-26.2a* possède un haplotype chez Norstar (N140), où une substitution dans le domaine AP2 (A>D) sur un acide aminé conservé et important pour la liaison à l'ADN, peut affecter la fonction de protéine. Le HC8 du domaine AP2 du gène *TaCBFIVd-4.1* chez le même cultivar est élargi à cause d'une substitution d'une Valine à la place de

Glutamate.

## CHAPITRE IV

### DISCUSSION

Ces dernières années, la famille des gènes *CBF* chez les *Triticeae* a été un sujet important de recherches. Le locus *Fr-2* s'est révélé important dans la régulation de la tolérance au gel chez le blé diploïde (Vágújfalvi et al. 2003), chez l'orge (Francia et al. 2004) et chez le blé hexaploïde (Toth et al. 2003; Båga et al. 2007). L'analyse des locus *Fr-2* a montré la présence de multiples gènes *CBF* chez le blé (Miller et al. 2006; Båga et al. 2007; Badawi et al. 2007) et chez l'orge (Francia et al. 2004; Skinner et al. 2006). Les analyses de séquences *CBF* chez les céréales suggèrent que les différentes espèces ont des familles *CBF* diversifiées et complexes. Nous avons maintenant identifié 65 gènes *CBFs* chez le blé hexaploïde représentant 27 groupes de gènes, avec un à trois homéologues de chaque gène. Ce travail représente une étape importante vers l'identification du répertoire complet des gènes *CBF* chez le blé hexaploïde. Les nombres élevés de gènes *CBFs* nous suggèrent qu'ils auraient des fonctions différentes. L'analyse HCA montre que basé sur la structure secondaire des protéines *CBF*, nous pouvons les présenter dans quelques dix-huit sous-groupes possédant des structures différentes, et qui pourraient affecter leur propriété et leur activité biologique. Comme déjà mentionné, plusieurs changements d'acides aminés situés dans les motifs/signatures *CBF*, le domaine AP2 et la région C-terminale ont été détectés dans notre étude. Ces changements ont, entre autres, comme effet de modifier la longueur des HC dans plusieurs domaines des protéines *CBF*. Les changements détectés dans le domaine AP2 peuvent avoir un effet sur l'interaction de ces protéines avec l'ADN. Les études ont montré que les gènes *CBF* démontrent des spécificités différentes pour l'élément cis lié (*HvCBFIVa-2* versus *HvCBFI-1*) (Xue, 2002 et 2003), et que des *CBF* appartenant aux groupes *CBFIV* aurait la propriété de devenir biologiquement actif seulement lors d'une baisse de température (Xue, 2003; Skinner et al. 2005). Il est donc probable qu'avec toutes les différences détectées au niveau de la structure du domaine AP2, certaines auront un impact sur la spécificité et/ou l'affinité de liaison à l'ADN et ainsi représenteront des groupes fonctionnels différents. Comme déjà mentionnés, nous avons trouvé des différences structurelles (longueurs/positionnement des noyaux hydrophobiques HC2, HC3, HC4, HC5 et HC6) au

niveau du domaine d'activation en C-terminal entre plusieurs groupes de protéines CBF. Ces différences suggèrent que ces groupes possèdent des repliements différents de cette région et ainsi des propriétés de transactivation différentes. Chez *Arabidopsis*, des changements effectués sur les noyaux hydrophobiques bordés par de courtes régions composées de résidus qui forment des boucles (Wang et al. 2005), peut modifier l'activité de transactivation, et ainsi la fonction de la protéine CBF. Malgré que la protéine CBF puisse tolérer la substitution d'acides aminés dans quelques positions, certains acides aminés hydrophobes sont importants dans le potentiel de régulation de transactivation de cette région (Wang et al. 2005). De plus, la protéine CBF peut interagir avec d'autres facteurs pour activer l'expression de gènes *COR*. Chez *Arabidopsis*, la protéine CBF1 active l'expression des gènes cibles en recrutant l'histone acetyltransferase (HAT), Gcn5 et des protéines adaptatrices de la transcription, Ada2 et Ada3 directement par son domaine AP2 (Mao et al. 2006). Le complexe contenant l'adaptateur HAT, remodèle la chromatine pour que le promoteur devienne plus accessible pour les facteurs de transcription (Stockinger et al. 2001). Il est donc possible que différents groupes CBF fonctionnels puissent être impliqués dans l'interaction avec des protéines adaptatrices spécifiques présentes seulement dans certains tissus ou stade développemental. De plus les groupes de CBF pourraient avoir des fonctions différentes qui se situent dans la température d'induction de leur expression à la basse température, ou l'expression spécifique tissulaire ou cellulaire des gènes *CBFs*.

Dans une étude précédente (Badawi et al. 2007), l'expression de plusieurs gènes des groupes IIIc, d et IVa, b, c, d a été associée avec le développement de la tolérance chez le cultivar de blé d'hiver plus tolérant au gel. En incluant les gènes identifiés dans notre étude, le blé hexaploïde aurait 20 groupes de gènes répartis dans les groupes IIIc, d, IVa, b, c et d qui pourraient être impliqués à rendre la plante plus résistante au gel. L'amplification spécifique des groupes IIIc, d, IVa, b, c et d chez les Pooideae s'est déroulée durant la période de refroidissement de l'ère Eocène-Oligocène (il y a environ 26 à 33.5 millions d'années) (Sandve et Fjellheim, 2010). La présence de plusieurs copies de gène protège la capacité du blé à développer une résistance au froid même dans le cas de la délétion d'une des copies. La présence

de gènes néo-fonctionnels aussi peut augmenter la tolérance du blé en améliorant sa capacité de résistance.

La présence de gènes *CBF* néo et sub fonctionnels dupliqués indique qu'ils sont toujours en évolution. La duplication des gènes rend le génome des organismes plus complexe (Ohno, 1970) et dans le cas des *CBF*, ceci peut augmenter la tolérance au froid en augmentant la régulation des gènes cibles (Jaglo-Ottosen et al. 1998; Jaglo et al. 2001; Liu et al. 1998).

Dans le cas où un organisme possède plusieurs copies d'un gène récemment dupliqué, la nature peut éliminer un des gènes redondants suite à l'accumulation de mutations délétères. Avoir des gènes sub-fonctionnels ou néo-fonctionnels grâce aux mutations ponctuelles qui changent les propriétés d'une des protéines ou l'expression d'un des gènes est une façon de garder toutes les copies d'un gène et ainsi augmenter le taux potentiel d'expression des *CBF*. Avoir plus d'une copie augmente les chances pour avoir un gène consensus qui fonctionnera à la suite d'une mutation ponctuelle (Matthew, 2009). L'étude phylogénétique montre qu'au courant de l'évolution des Pooideae, le nombre total de gènes *CBF* a été augmenté, contribuant ainsi à l'augmentation de la tolérance au gel du blé.

Comme les gènes *CBF* sont associés avec la tolérance au gel et le locus *FR-2* (Stockinger et al. 2007; Vágújfalvi et al. 2005), ils représentent des candidats appropriés pour ce QTL contrôlant la tolérance au gel chez les *Triticeae*. Donc dans une deuxième étape, nous avons essayé de trouver des polymorphismes dans la région codante des gènes *CBF* qui pourraient causer des modifications importantes d'acides aminés et ainsi influencer la structure des protéines *CBF* et leur fonction dans le développement de la tolérance au gel. Premièrement, notre analyse montre que Norstar possède 2 pseudogènes *CBF* exprimés et 24 des 64 protéines analysées ont une séquence/structure qui dévie de celle conservée dans les autres membres apparentés du groupe. Il est possible de prévoir que certains de ces changements auront une influence négative sur l'activité biologique de ces protéines, et ceci aurait une répercussion sur le développement de la résistance maximale au gel. Ces gènes pourraient donc être considérés

comme des allèles moins optimales (ou sub optimales). Alternativement, il est possible que certaines de ces protéines dont la structure/séquence dévie, pourraient avoir un effet positif et que les gènes correspondant auraient été maintenu par la pression de sélection. Des travaux futurs sur l'activité de ces variantes pourront trancher entre ces possibilités. De plus, nos résultats montrent que les cultivars de printemps, Chinese Spring et Manitou, possèdent en grande partie les mêmes gènes non fonctionnels/sub optimaux qui existent chez Norstar. Chez ces cultivars, les gènes du groupe *CBFIIIc-13* et *CBFIVd-B22* identifiés sont non fonctionnels. Donc même si ces gènes s'exprimeraient à basse température, leurs protéines ne pourraient pas activer les gènes *COR* cibles et ainsi induire la tolérance au gel. Par contre, le gène *CBFIIIc-13* chez le blé diploïde n'est pas un pseudogène (Knox et al. 2008). Le blé hexaploïde n'est pas le seul à posséder des pseudogènes. Le gène *CBFIIIc-12* chez le blé diploïde (Knox et al. 2008), et les gènes *CBFIIIc-13* (cultivar Tremois) (Knox et al. 2010) et *CBFIIIc-8* (Skinner et al. 2005) chez l'orge, ne sont pas fonctionnels. En effet le gène *CBFIIIc-13* contient plusieurs mutations non-sens chez le cultivar d'orge plus sensible au gel (Knox et al. 2010). De plus, chez le cultivar de blé diploïde sensible, la protéine du gène *CBFIIIc-12* est non fonctionnelle du fait de la délétion de cinq acides aminés dans le domaine de liaison à l'ADN (Knox et al. 2008). La délétion de nucléotides dans la région codante des trois gènes de l'orge, *HvCBF8A*, *HvCBF8B* et *HvCBF8C*, les transforme en pseudogènes chez les cultivars sensibles et résistants au froid (Skinner et al. 2005). Ainsi une des caractéristiques des cultivars sensibles est la présence d'un nombre élevé de pseudogènes avec des protéines non fonctionnelles. Chez le blé hexaploïde, notre résultat indique qu'une différence dans le contenu de pseudogènes ne serait pas suffisante pour expliquer la tolérance au gel plus élevée chez le cultivar d'hiver. Toutefois, il est toujours possible que parmi les gènes *CBF* du répertoire qui n'ont pas été ré-amplifié, il ait une différence dans le contenu de pseudogènes entre les cultivars de printemps et d'hiver.

L'analyse des haplotypes retrouvés chez les deux cultivars de printemps montre que le gène *CBFIIIc-24.2b* semble être meilleur/plus optimal en comparaison avec l'haplotype chez Norstar tandis que l'haplotype *CBFIVb-A20b* semble être moins optimal que la variante chez Norstar. L'analyse des haplotypes présents spécifiquement chez un cultivar a aussi permis



quelques observations intéressantes. Chez Norstar, l'homéologue *TaCBFIVd-4.1* a un changement qui rend la région hydrophobique HC2 plus grande dans la région C-terminale. Chez les cultivars de printemps Chinese Spring et Manitou, *TaCBFIVd-4.1* possède des substitutions d'acides aminés additionnelles suggérant qu'elles pourraient représenter des versions encore moins optimales. Les deux variantes de *TaCBFIVd-B4* détectés chez Chinese Spring sont non-fonctionnelles à cause de l'insertion d'une base. Le seul variant de *TaCBFIVd-D9* détectés chez Manitou est non fonctionnel à cause d'une mutation non-sens, et chez Chinese Spring le variant S275 possède une substitution d'un acide aminé dans une position très conservée et importante pour la liaison à l'ADN. Il est donc possible qu'à des niveaux d'expression similaires, l'activité biologique des protéines du groupe IVd est plus faible chez les cultivars de printemps. Comme les membres de ce groupe sont exprimés plus fortement chez le cultivar d'hiver (Badawi et al. 2007), ceci peut amplifier l'inefficacité de ce groupe à promouvoir le développement de la tolérance chez les cultivars de printemps.

L'analyse de toutes les copies de *TaCBFIVc-14.1* retrouvés chez Manitou semblent indiquer qu'ils sont non- ou moins-fonctionnelles à cause d'une substitution d'un acide aminé dans une position très conservée dans la signature CBF. Canella et ses collègues (2010) ont montré que la délétion ou des mutations au sein de la signature CBF, en avant du domaine AP2 peut considérablement nuire à la capacité de CBF1 d'*Arabidopsis* d'induire l'expression des gènes *COR* cibles. Cette déficience fonctionnelle n'était pas due à des mutations entraînant une instabilité des protéines ou de la perte de l'importation de protéines dans le noyau. Plutôt, ces mutations ont considérablement nui à la capacité de CBF1 de se lier au motif CRT/DRE. Donc basé sur cette étude (Canella et al. 2010), il est raisonnable de prévoir que les protéines des haplotypes mutés de *TaCBFIVc-14.1* chez Manitou auront une déficience fonctionnelle pour se lier sur l'ADN de gènes *COR* cibles. Le HC4 de la région C-terminale du gène *CBFIVc-B14* est plus long et peut modifier la structure secondaire de la protéine et son interaction avec des protéines adaptatrices de transcription. Les deux substitutions dans le HC3 du domaine C-terminal de *CBFIVc-14.3*, semblent aussi le rendre moins optimal. Ces variants de *CBFIVc-B14* et *CBFIVc-14.3* sont présents chez les trois cultivars indiquant que le seul gène consensus du

groupe IVc est le *14.1* et Manitou possède un gène *TaCBFIVc-14.1* moins optimal pour aider la plante à résister aux conditions de gel. Comme les gènes du groupe IVc sont exprimés 10 fois plus chez Norstar comparativement aux cultivars de printemps (Badawi et al. 2007), et que le nombre de gènes dans le groupe IVc est petit (un seul), ce groupe peut être sensible à des mutations qui affecteraient l'activité biologique d'un de ces homéologues (copies). La présence de plus de gènes optimaux dans ce sous-groupe pourrait aider à mieux induire les mécanismes de tolérance et ainsi survivre au gel. Les substitutions des acides aminés impliqués dans l'interaction avec d'autres facteurs, affecteraient le potentiel de trans-activation des protéines CBF. Une diminution de l'activité de trans-activation des CBF, diminuerait l'expression de gènes *COR* et le résultat serait un blé moins résistant au gel. D'autres gènes comme *CBFIIIId-19* et *CBFIVa-2* ont montré des expressions 100 fois plus élevées chez le cultivar d'hiver (Badawi et al. 2007) et existeraient sous forme de plusieurs variantes chez le blé. Comme ces différences se trouvent dans des positions très conservées, il est possible d'envisager que l'expression inégale chez ces cultivars cause des effets sur le développement de la tolérance. De plus, des différences polymorphiques entre les deux cultivars peuvent être localisées sur le promoteur de ces gènes.

Une autre différence entre les cultivars résistants et sensibles au froid est la présence de plusieurs copies/paralogues qui augmente la dose de gènes fonctionnels et qui est un point fort pour développer la résistance au gel. Le cultivar Nure de l'orge qui est résistant au froid contient deux copies de *CBF2* (*CBF2A* et *CBF2B*) et deux copies du gène *CBF4* tandis que Tremois qui est un cultivar sensible au froid possède une copie de *CBF2* et une copie de *CBF4* (Knox et al. 2010). Le gène *CBF14* est exprimé à un niveau plus élevé chez le blé tolérant au gel par rapport au cultivar moins tolérant (Vágújfalvi et al. 2005). Il semble que le nombre de copies de ce gène chez le cultivar résistant est augmenté (Knox et al. 2010). Comme suite à notre étude, il serait souhaitable de séquencer les régions génomiques entourant les gènes *CBF* identifiés chez le blé afin de vérifier l'existence de copies de gènes *CBF* et pour trouver s'il y aurait une association entre le nombre de copies d'un gène *CBF* dans un génome et la capacité du cultivar à développer la tolérance au gel.

Dans cette étude, nous avons identifié des gènes *CBF* nouveaux chez le blé hexaploïde Norstar. Les protéines *CBF* du blé peuvent être subdivisées en 18 groupes démontrant des différences structurales dans les régions AP2 et C-terminale. Ces différences suggèrent que les groupes auront des propriétés/fonctions différentes. Nous avons ensuite essayé d'identifier les différences dans les régions codantes entre les gènes *CBFs* chez un cultivar d'hiver (Norstar) et deux cultivars de printemps (Chinese Spring et Manitou). Le but était d'identifier les causes possibles du comportement différent du blé d'hiver et du blé de printemps au gel. Notre analyse a montré que le cultivar d'hiver ne contient pas juste des allèles optimaux qui puissent aider la plante à développer une résistance au gel maximale. De plus, l'analyse des cultivars de printemps montre qu'ils possèdent des allèles plus ou moins optimaux que chez Norstar. Comme le gène *TaCBFIIIId-24.2a* qui semble être le meilleur, mais *TaCBFIVb-A20b* qui est probablement moins optimal. La différence fonctionnelle de ces protéines peut modifier leur capacité à interagir avec les gènes *COR* et par conséquent influencer la réponse de la plante au gel. Il y a des gènes avec plus d'une copie, il est possible que la redondance de gènes offre une protection contre la délétion de ce gène dans un génome ou augmente la dose d'un gène fonctionnel important. La présence d'un grand nombre de groupes fonctionnels de gènes *CBFs*, combiné avec possiblement des spécificités d'expression, peut permettre à la plante une meilleure acclimatation au froid afin de développer une tolérance au gel plus forte.

La diversité moléculaire détectée dans les gènes provenant de cultivars différents est représentative de celle qui existe dans les populations de blé et nous indique une grande variation génétique au fil du temps (Dubreuil et al. 1996; Rebourg et al. 1999). La mutation est le point de départ du polymorphisme et de l'évolution moléculaire. L'analyse des polymorphismes présents chez les cultivars testés montre que la famille *CBF* est composée de plusieurs catégories de gènes : pseudogènes (codant pour des protéines non fonctionnelles), des gènes optimaux (protéines démontrant des structures/séquences identiques à la séquence consensus), et des gènes sub-optimaux (protéines démontrant des structures/séquences déviant de la séquence consensus). De plus, cette analyse laisse entrevoir que les différents cultivars auront leur propre répertoire unique de ces différentes catégories de gènes. A court terme, il sera important d'étendre ce type

d'analyse à d'autres cultivars et de vérifier si les prédictions théoriques décrivant les allèles optimaux et les allèles sub-optimaux se traduisent effectivement par des changements de l'activité biologique des protéines correspondantes qui pourraient influencer la régulation des gènes *COR* et ainsi le développement de la tolérance au gel. Les allèles optimaux retrouvés chez différents cultivars représenteraient des marqueurs moléculaires potentiels de résistance au gel, et pourraient donc devenir des candidats intéressants dans le cadre de programmes de sélection assistés qui tenteraient de réunir dans un cultivar donné, une majorité d'allèles CBF optimaux. De plus, des recherches futures viseront : l'étude approfondie de l'expression des gènes *CBFs* pendant l'acclimatation au froid de cultivars résistant et sensible au gel; et la recherche chez ces cultivars de différences moléculaires dans les régions régulatrices des gènes *CBFs* qui pourraient être associées avec une expression plus élevée de gènes CBF. Ces travaux permettront d'identifier les gènes associés au processus d'acclimatation et possiblement des candidats additionnels de marqueurs moléculaires potentiels de résistance au gel qui pourraient ajouter aux choix disponibles pour les sélectionneurs de plantes.

**ANNEXE****LISTE DES FIGURES ET TABLEAUX SUPPLÉMENTAIRES**

Figure	Page
S1 Liste des séquences codantes de CBF chez <i>T. aestivum</i> .....	81
S2 Liste des séquences codantes de CBF chez <i>T. monococcum</i> , <i>H. vulgare</i> et <i>S. cereale</i> .....	99
S3 Liste des séquences codantes de cultivars Norstar, Chinese Spring et Manitou (de <i>T.aestivum</i> ) démontrant des polymorphismes spécifiques.....	121

Tableau	Page
S16 Amorces spécifiques utilisées pour ajouter des informations sur la séquence des gènes TaCBF. ....	136
S27 Amorces spécifiques utilisées pour amplifier les nouveaux gènes TaCBF.....	137
S38 Amorces spécifiques utilisées pour amplifier les gènes TaCBF à partir de l'ADN génomique de différents cultivars de blé. ....	139



**Figure S1: Liste des séquences codantes de CBF chez *T. aestivum*.**

Les régions surlignées en jaune/rouge sont les sequences signature/AP2 utilisées dans les analyses phylogénétiques. Les gènes et les numéros d'accèsion mis en évidence en rose ont été utilisés pour produire les figures 1 à 3. Toutes les séquences ont été traduites et analysées par HCA.

**TaCBFI-A11 (EF028751)**

ATG GAGTGGGCGTACAGCGGCGGCCACTCGTCGTCGGGAACGAAATCTCCGGCGGCCGGCGGCCGGGAGGAGGG  
CTCGTACATGACGGTGTCTCGGCGCCG **CCCAAGCGGCGGGCGGGAGGACCAAGGTCAGGGAGACGAGGCACCCGG**  
**TGTACAAGGGGGTGCAGCAGCAACCCCGGCGGTGGGTCTGCGAGGTGCGCGAGCCGACGGGAACAGCGGGCTG**  
**TGGCTCGGCACCTTCGACACCGCGGAGATGGCGGCGCGCGCACGACGTGCGCGCCCTCGCGCTCCCGCGCCGCGC**  
**AGCGTGCCTCAACTTCGCG** **GACTCGCGCGCACG**CTCCGGGTGCCGCCGAGGGGGCCGGCCACGAGGAGATACGCC  
GCGCCGCGGTTCGAGGCGGCCGAGCTGTTCGCCCGGAGCCTGGGCAGCGCAATGCAGCTACTACCGAGGCGCCGGCT  
GCTTCACCGGCAGACGCGGGGAACGCGGAGCTCGTTGCAACTCTCCTTACCACCTCATGGATGGTTAGAAATTCGA  
AATGCAGGGCTATCTTGACATGGCGCACGGCATGCTGATCAGGCCACCGCCAATGGCGGGGCGCTGCAGCTGGATCG  
AGGAGGACTACGACTGCGAGGTACGCTGTGGAAGTAC **TGA**

**TaCBFI-1.1 (Accession)**

ATG GACCTCAGTGCCCTCAGCAGCGACTACTCGTCGGGGACGCCGTCCCCGGTGGGCGCGGACGGCGGCAACAGCGA  
GGGCTTCTCGACGTACATGACGGTGTCTTCGGCTCCG **CCGAAGCGCGCGCGGGCGGACCAAGTTCAAGGAGACGC**  
**GGCACCCTG**CTCTACAAGGGCGTGCGCCGAGGAACCCCGGAGGTGGGTCTGCGAGGTGCGGGAGCCGACAGCAAG  
**AGAGGATATGGCTCGGCACGTTTGAGACCGCAGAGATGGCGGCACGCGCGCACGACGTGGCCGCGCTGGCGCTGCG**  
**TGGCGCGCGCGCTGCTGAACCTCGCG** **GACTCGCCTCGCGCG**CTCCGGGTCCCGCCCGTGGGTGCTAGCCCTGATG  
AGATACGGCGGGCGCGGTAGAGGCGGCTGAGGCATTCTGCCGGCACCCGTCCAGAGCAATGCGGCCGCCGAGGAG  
GCAGCCGTTGCACCAACGGAGCAGTTCGCCGGTGACCCGTACTATGGAATGGACGATGGGATGGACTTCGGGATGCA  
GGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCTCCGATGGCGGGCCCGTCAGCGACTGTGCGGAGAGG  
GCGACGATGACGGCGAGGTACGCTTGTGGAGCTAC **TGA**

**TaCBFI-1.2 (Accession)**

ATG GACCTCAGTGCCCTCAGCAGCGACTACTCGTCGGGGACGCCGTCCCCGGTGGGCGCGGACGGCGGCAACAGCGA  
GGGCTTCTCGACGTACATGACGGTGTCTTCGGCTCCG **CCGAAGCGCGCGCGGGCGGACCAAGTTCAAGGAGACGC**  
**GGCACCCTG**CTCTACAAGGGCGTGCGCCGAGGAACCCCGGAGGTGGGTCTGCGAGGTGCGGGAGCCGACAGCAAG  
**AGAGGATATGGCTCGGCACGTTTGAGACCGCAGAGATGGCGGCACGCGCGCACGACGTGGCCGCGCTGGCGCTGCG**  
**TGGCGCGCGCGCTGCTGAACCTCGCG** **GACTCGCCTCGCGCG**CTCCGGGTCCCGCCCGTGGGTGCTAGCCCTGATG  
AAATACGGCGGGCGGCGGTAGAGGCGGCTGAGGCATTCTGCCGGCACCCGACCAGAGCAATGCGGCCGCCGAGGAG  
GCAGCCGTTGCACCATCGGAGCAGTTCGCCGGTGATCCCTACTATGGGATGGACGATGGGATGGACTTCGGGATGCA  
GGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCTCCGATGGCAGGCCCGTCGGCGACTGTGCGGAGACG  
GCGACGATGACGGCGAGGTACGCTTGTGGAGCTAC **TGA**

**TaCBFI-1.3 (Accession)**

ATG GACCTCAGTGCCCTCAGCAGCGACTACTCGTCAGGACGCCGTCCCCGGTGGGCGCGGACGGCGGCAACAGCGA  
GGGCTTCTCGGCGTACATGACGGTGTCTTCGGCGCCG **CCGAAGCGCGCGCGGGCGGACCAAGTTCAAGGAGACGC**  
**GGCACCCTG**CTCTACAAGGGCGTGCGCCGAGGAACCCCGGAGGTGGGTCTGCGAGGTGCGGGAGCCGACAGCAAG  
**AGAGGATATGGCTCGGCACGTTTGAGACCGCGGAGATGGCGGCACGCGCTCACGACGTGGCCGCGCTGGCGCTGCG**  
**TGGCGCGCGCGCTGCTGAACCTCGCG** **GACTCGCCTCGCGCG**CTCCGGGTCCCGCCCGTGGGTGCTAGCCCTGATG  
AGATACGGCGGGCGGCGGTGAGGCGGCTGAGGCATTCTGCCGGCACCCGACCAGAGCAATGCGGCCGCCGAGGAG  
GTAGCAGTTGCACCAACGGAGCAGTTCGCCGGTGATCCGTACTATGGGATGGACGATGGGATGGACTTCGGGATGCA  
GGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCTCCGATGGCAGGCCCTTCAGCGACTGTGCGGAGACG  
GCGACGATGACGGCGAGGTACGCTTGTGGAGCTAC **TGA**



## &gt;TaCBFII-5.1a(EF028752)

ATG GACCAGTACAGCTACCGCGGCGGGGACGACAATGGCCAGGGCGGGTACGCGACGGTGACGTCGGCGCCGCC  
 GAAGCGGCGGGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCG GTGTACCGGGGCGTGCGCCGGCGCGCGCGG  
 TCGGGCGGGTGGGTCTGCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTCGCCTCCCCCGAG  
 GCGCGGGCCCGCGCCACGACGTTGCCGCGCTGGCCCTGCGCGGCGCGCGCGGTGCCTCAACTTCGCG GACTCGGC  
 GGCCCTGCTCGCCGTCGACCCCGCCACGCTCCGCACCCCGACGACATCCGCGCGGCGGCCATCGCGCTCGCCGAGA  
 CTGCCTGCCCCGCCGCGCGGCCTCGTCGTGCGCCGTGGCCGCCGTGGCGTCCGCGCGGCGCCCCGATGATGACG  
 ATGATGCACGAGTCCGCGGCGGTGCACTACGACGACTACCCGATGCAGTACGGGTACGGCGGCATTGGGGACTTGGA  
 CCAGGATTCTACTACTACGACGGGATGAGCGCCGCCGGCGGCGACTGGCAGAGCGGCTCGCACATGGACGGAGCCG  
 ACGACGACTGCAACGACAGCGGTGGGTACGGCGCGGCGAGGTCCCGCTCTGGAGCTAC TGA

## &gt;TaCBFII-5.1b(Accession)

ATG GACCAGTACAGCTACCGCGGCGGGGACGACAATGGCCAGGGCGGGTACGCGACGGTGACGTCGGCGCCGCC  
 GAAGCGGCGGGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCG GTGTACCGGGGCGTGCGCCGGCGCGCGCGG  
 TCGGGCGGGTGGGTCTGCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTCGCCTCCCCCGAG  
 GCGCGGGCCCGCGCCACGACGTTGCCGCGCTGGCCCTGCGCGGCGCGCGCGGTGCCTCAACTTCGCG GACTCGGC  
 GGCCCTGCTCGCCGTCGACCCCGCCACGCTCCGCACCCCGACGACATCCGCGCGGCGGCCATCGCGCTCGCCGAGA  
 CTGCCTGCCCCGCCGCGCGGCCTCGTCGTGCGCCGTGGCCGCCGTGGCGTCCGCGCGGCGCCCCGATGATGACG  
 ATGATGCACGAGTCCGCGGCGGTGCACTACGACGACTACCCGATGCAGTACGGGTACGGCGGCATTGGGGACTTGGA  
 CCAGGATTCTACTACTACGACGGGATGAGCGCCGCCGGCGGCGACTGGCAGAGCGGCTCGCACATGGACGGAGCCG  
 ACGACGACTGCAACGCGAGCGGTGGGTACGGCGCGGCGAGGTCCCGCTCTGGAGCTAC TGA

## &gt;TaCBFII-5.2(EF028753)

ATG GACCAGTACAGCTACCGCGGCGGCGACGCGGACCAGGCCGGGCGGGTACGCGACGGTGACGTCGGCGCC  
 GCCCAAGCGGCGGGCGGGGCGGACCAAGTTCCGGGAGACGCGCCACCCG GTGTACCGGGGCGTGCGCCGGCGCGCGG  
 TCGGGCGGGTGGGTCTGCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTCGCCAGCCCC  
 GAGGCGCGCGCGCGCGCCACGACGTCGCCGCGCTGGCCCTGCGCGGCGCGCGCGCTGCCTCAACTTCGCG GACTC  
 GGCCACGCTGCTCGCCGTCGACCCCGCCACGCTCCGCACCCCGACGACATCCGCGCGCGCGCCATCGCGCTCGCCG  
 AGACCGCCTGCCCCGCCGCGCCCGTGGCCGCGGAGGCGTCCGCGCGGCGCGCCGCGATGATGGCGATGATGCAAGAG  
 CCCTCGGCGGTGGAGTATGACGACTACCCGATGCAGTACGGCGGCATTGGGGACTTCGACCAGCATTTCGTACTACTA  
 CGACGGGTTGAGCGCGGCGGCGGCGACTGGCAGAGCAGCTCGCACATGGACGGAGCCGACGACGACTCAAACCTGCG  
 GCGGTGGGTACGGCGCCGCGGAGGTCCCACTTTGGAGCTAC TGA

## &gt;TaCBFII-5.3(EF028754)

ATG GACCAGTACAACTACGGCGGCGGCGTGGCCTACTACGGCAGCACCACCGCGGCGTAGGCGACAACGGCCAGGG  
 CGGCGGGTACGCGACGGTGACGTCGGCGCCG CCGAAGCGGCGGCGGGGCGGACCAAGTTCCGGGAGACGCGCCACC  
 CGGTGTACCGGGGCGTGCGCCGGCGCGGCGCGGCGGGCGGTGGGTGTGCGAGGTGCGCGAGCCCAACAAGAAGTCC  
 UGCATCTGGCTCGGCACCTTCGCCAGCCCCGAGGCGCGGCGCGCCACGACGTCGCCGCGCTCGCGCTCCGGGG  
 CCGCGCGGCTGCCTCAACTTCGCG GACTCGGCCACGCTGCTCGCCGTCGACCCCGCCACGCTCCGCACCCCGACG  
 ACATCCGCGCGCGCCGATCGCGCTCGCGGAGACCGCTGCCCGCGCGCGCGCGTCTCTCGTCCGTGGCCGCC  
 GCGGTGGCGTCCGCGCGGCGGCCCCGATGACGATGATGCAGTTTGACGACTACGCGATGCAGTACGGCGGCATTGG  
 GGACTTGGACCAGCATTCCTACTACTACGACGGGTTGAGCGCCGCGGCGGCGACTGGCCGAGCGGCTCGCACATGG  
 ACGGAGCCGACGACGACTGCAACGCGACGCGGTGGCTACGGCGCGGCGAGGTGCGGCTCTGGAGCTAC TGA

## &gt;TaCBFII-5.4(Accession)

ATG GACAACTCCGCGTGATCTTCTATGGCGGCGCATACGCGACGGTGATGTCGGCGCCG CCGAAGCGGCGGCGGG  
 GCGGACCAAGTTCCGGGAGACGCGCCACCCG GTGTACCGGGGCGTGCGCCGGCGCGGCGCGGCGGGCGGTGGGTGT  
 GCGAGGTGCGCCAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTCGCCAGCCCCGAGGCGCGCGCGCGCG  
 TACGACGTCGCCGCGCTCGCGCTCCCGGGGCGCGCGCGCTGCCTCAACTTCGCG GACTCCGCGCGCTGCTCGCCGT  
 CGACCCGGCCACGCTCCGCACGCCCCAGGACATCCGCGCGCGCGCAATCAGCTCGCCCGGGCGCCTGCCCGCACG  
 ACGCGACGAGGTCTCTGTGTCCGCGGCGTCTGCGCCGGCGCGCGCGATGGTGATCAGCAGGAGCGCGGCTACG  
 CCGTACGACAGCTACGCCATGTACGCGGCGTTGGCGGACCTGGAACAGCATTCCCACTGCTACTACGACGGGATGAG



CGGCAGCGGCGACTGGCAGAGCATCTCACACATGAACGTCGCCGACGAAGACGGTGGCTACGGCGCAGGAGACGT  
CGCGCTCTGGAGCTAC**TGA**

>TaCBFII-5.5 (Accession)

**ATG**GAACAGTGC GGCGTGGGCCTCTACGGCGTCTGTCGAGGGCAGCGGATACGCGACGGTGACTACCGCGCCG**CCTAA**  
GCGGCCGGCGGGGGCGGACCAAGTTCCGGGAGACGCGCCACCCG**CTCTACCGCGCGGTGCGCCGGCGCGGGCGCGCGG**  
GCGGGTGGGTGTGCGAGGTGCGCCAGCCCAACAAGAAGTTCGCGCATCTGGCTCGGCACCTTCGCCACGCCCGAGGCC  
GCCGCGCGCGCCACGACGTGCGCGCTCGCGCTCCGGGGCCGAGCCGCCTGCC**TCAACTTCGCCGACTCGGCCAC**  
**GCTT**CTCGCCGTGACCCCGCCACGCTCCGCACGCCCCAGGACATCCGCGCCGCCGAATCGCGCTCGCCCAGGCGG  
CCTGCCCGCACGACGCGAGGAGGTCTCTGTGTCCGTGGCGTCCGCGCGGGCGCCCGCGATGGTGATCATGGAGGAG  
GCCGCGGCGGCACCGTACGACAGCTACGCCATGTACGGCGGCTTGGCGGACCTGGACCAGCATTCTACTGTCTACTC  
CAACGGGATGAGCGGCGGCGGCGACTGGCAGAGTATCTCGCATATGGACGGAGCCGACGAAGACGGCAGCTACGGCG  
CAGGAGACGTGCGCTCTGGAGCTAC**TGGT**CGCGTGGGATCGATCGGGCAGATTGT**TGA**

>TaCBFIIIa-6.1 (EF028755)

**ATG**TGTCCGATCAAGAAGGAGATGAGCGGGGAGTCCGGCTCGCCGTGCAGCGGGGAGAACTTCTACTCGCCCTCCAC  
GTCGCGGGAGCATCAGCAGGCGAAGCAGGCGGCGTGGACGTGCGCGCCG**GCGAAGCGGCCGGCGGGGGCGGACCAAGT**  
**TCAGGGAGACGCGGCACCCG**GTATACCGCGCGTGC GGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGC  
**GTGCCCCGGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGACACCGCCGAGGCCCGCGCGCGCCCAACGACGC**  
**CGCCATGATCGCGCTCTCCGCGGGGGCGCGGGCTGCCTCAACTTCGCCGACTCCGCGGAGCTGCTCGCCGTGCCGG**  
CGGCTCCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTCTGGAGGCCGTCGAGGACTTCCTCGCGCGGGAG  
CGCATCGCCGAGGAGACGCGCTCTCGGGACCTCGTCTCGCGCCCTCCTCCCTACCGACGACGAGTCTGCTCCTC  
TCCCCGCCCGAGGACTCGCCGTTTCGAGCTGGATGTCCTGAGCGACATGGGCTGGGACCTGTACTACGCGAGCCTTG  
CGCAGGCGATGCTGATGGCGCCGCTTCTTCCATGGCCGCGGCGCTCGGCGATTACGGCGAGGTGATGTGCCACTC  
TGGAGCTACCAGAGC**TAG**

>TaCBFIIIa-6.2 (EF028756)

**ATG**TGTCCGATCAAGAGGGAGATGAGCGGGGAGTCCGGCTCGCCGTGCAGCGGGGAGAACTTCTGCTCGCC  
CTCGGCGTTCGCCGAGCGCCAGCAGGCGAGGCGGGGTGGACGTGCGCGCCG**GCGAAGCGGCCGGCGGGGGCGGA**  
**CCAAGTTTCAGGGAGACGCGGCACCCG**GTGTACCGCGCGTGC GGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
**GTGCGCGTGC CGCGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGACACCGCCGAGGCCCGCGCGCGCCAA**  
**CGACGCCGTCTGCTCTGCTCGCCGCCGAGGCGCGCCTGCCTCAACTTCGCCGACTCGGCCGAGCTGCTCTCCG**  
TGCCGTTGGCTCCTCCTACCGCAGCCTGGACGAGGTCCGCCACGCCGTCTGGAGGCCGTCGAGGACTTGCTGCGG  
CGCGAGGCGCTCGCCGAGGAGGACGCGCTCTCGGGACCTCCTCGTCCGCGCCCTCCCCCTACCGACGACGAGTC  
GTCTCTTCGCCGTGCCGAGGAGACTCGCCGTTTCGAGCAGGACGTGCTGAGCGAGATGGGCTGGGACCTGTACT  
ACGCGAGCCTGGCGCAGGCGATGCTCATGGCGCCGCCCGCGGCTGCGGCGCTCGGCGATTACGGGGAGGCTCAC  
CTCGCCGACGTGCCACTCTGGAGCTACCAGAGC**TAG**

>TaCBFIIIa-D6 (EF028757)

**ATG**TGTCCGATCAAGAGGGAGATGAGCGGGGAGTCCGGCTCGCCGTGCAGCGGGGAGAGCTTCTACTCGCCCTCCAC  
GTCGCCGAGAACAGCAGGCGAGGCGGCGTGGACGTGCGCGCCG**GCGAAGCGGCCGGCGGGGGCGGACCAAGT**  
**TCAGGGAGACGCGGCACCCG**GTGTACCGCGCGTGC GGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGC  
**GTGCCCCGGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGACACCGCCGAGGCCCGCGCGCGCCAAACGACGC**  
**CGCCATGATCGCGCTCTCCGCGGGGGCGCGGGCTGCCTCAACTTCGCCGACTCGGCCGAGCTGCTCGCCGTGCCGG**  
CAGCTCCTCCTACCGCAGCCTCGACGAGGTCCGCCACGCCGTCTGGAGGCCGTCGAGGACTTGCTGCGGCGCGAG  
GCGATCGCCGAGGACGACGCGCTCTCCGGACCTCCTCGTCCGCGCCCTCCTCCCTACCGACGACGGGTGCTCCTC  
CTCGCCGCTGCCGAGGAGGACTCGCCGTTTCGAGCTGGACGTGCTGAGCGACATGGGCTGGGACCTGTACTACGCGA  
GCCTGGCGCAGGCGATGCTCATGGCGCCGCTTCTTCCATGGCCGCGGCGCTCGGCGATTACGGCGAGGTGATGTG  
CCTCTGAGCTACCAGAGC**TAG**

>TaCBFIIb-23.1 (Accession)



ATG GACATGAGCCTCGAGCACTCGAGCTCTCCCTCCTCCTCATCCACTACCGAGCGCGCCGGGACGGCGTGCCG  
TGGCCGCGGAAGCGCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCCGTGTTCGGCGGGCGTGCGCGCGCG  
TGGAGCGCGCGCGCGTGGGTCTGCGAGGTGCGCGTGCCTCGGGGAGCGCGGCACCGCGCTCTGGCTCGGGACGTACA  
TCAATGCGCGAGCGCGCGCGCGAGCGCACGACGCGCGCATGCTCATGCTCCGGGGCGGCTCCGCGCGGTGCGCTCAAC  
TTCGGGACTCCGCGTGCTGCTCGCGGTGCCGCCGCGATTACCAACCTCTCTGACGTCCGGCGCGCGCGCGCTCCA  
GCGCGTGCGGGACTTCTGCGCCGTCCGGAGGCCAACAGTGCCGCTGCCGCGGTGCAGGAGGTACCTCCAGCGTGT  
GCGCCCGCTCGTCCGCGGTGTGAGTGTGCCCTCGTCAGAGACGGCGCAGGCTTCCCTGATGCCGCTTTTGAAGCG  
GCGCCCGCACTGGACATGGACGTGTTTGAACCTCGACTGCTTGTGTTGGGAAACGGACTCCGACGCGTACTACTACGC  
TACCTTGCGCAGGGGCTGCTCATGGAGCCACCGCCGACTGTGCCACCGGGTCTACTGGGACAATGGAGACTGCG  
GCGCGCGGAGCCGGAGCTGATGTGCGCTCTGGAGTTAC **TAC**

**TaCBFIIIb-23.2 (Accession)**

ATG GACATGAGCCTCGAGCACTCGAGCTCTCCCTCCTCCTCATCCAGACCGAGCGCGCCGGAACGGCGTGCCGCTG  
CGCGCGGAAGCGCCCGCGTGGCCGCACTAAGTTCCGGGAGACGCGGCACCCCGTGTTCGGCGGGCGTGCGCGCGCGT  
GTAATGCGCGCGCGTGGGTCTGCGAGGTGCGCGTGCCTCGGGGAGCGCGGCACCGCGCTCTGGCTCGGGACGTACATG  
ACCAATGCGAGCGCGCGCGCGCGCGCACACGACGCGCGCATGCTCATGCTCCGGGGCGGCTCCGCGCGGTGCGCTCAACT  
GTAAGACTCCGCGTGCTGCTCGCGGTGCCGCCGCGTTACCAACCTCTCTGACGTCCGGCGCGCGCGCGCTCCAGG  
CGGTGCGGGACTTCTGCGCCGTCCGGAGGCCAACAGTGCCGCTGCCGCGGTGCAGGAGGTACCTCCAGCGTGTCC  
GCGCCGCTCGTCCGCGAGTGTGTAGTGTGCCCTCGTCAGAGACGGCGCAGGCTTCCGCTGATGCCGCTTTTGAAGCGCG  
GCGCGCACTGGACATGGACGTGTTGACCTTGACTGCTTGTGTTGGGAAACGGATTCCGATGCGTACTACTACGCGA  
TCTTGCGCAGGGGCTGCTCATGGAGCCACCGCCGACTCTGCCACCGGGTCTACTGGGACAATGGAGACTGCGCC  
TTCGGCGAGCCGGAGCTGATGTGCGCTCTGGAGTTAC **TAC**

**TaCBFIIIb-23.3 (Accession)**

ATG GACATGAGCCTCGAGCACTCGAGCTCTCCCTCCTCCTCATCCAGACCGAGCGCGCCGGGACGGCGTGCCGCTG  
CGCGCGGAAGCGCCCGCGCGCGCACCAAGTTCCGGGAGACGCGGCACCCCGTGTTCGGCGGGCGTGCGCGCGCGT  
GTAATGCGCGCGCGTGGGTCTGCGAGGTGCGCGTGCCTCGGGGAGCGCGGCACCGCGCTCTGGCTCGGGACGTACATGAC  
GAGGAGCGCGCGCGCGCGCGCACACGACGCGCGCATGCTCATGCTCCGGGGCGGCTCCGCGCGGTGCGCTCAACTTCCG  
GACTCCGCGTGCTGCTCGCGGTGCCGCCGCGTTACCAACCTCTCTGACGTCCGGCGCGCGCGCGCTCCAGGCCG  
GCGGACTTACTGCGCCGTCCGGAGGCCAACAGTGCCGCTGCCGCGGTGCAAGAGGTACCTCCAGCGTGTCCGCC  
GCTCTCGCGGTGTGCAGTGTGCCCTGTGTCAGAGACGGCGCAGGCTTCCGCTGATGCCGCTTTTGAAGCGCGCGG  
TGCACTGGACATGGACGTGTTGACCTCGACTGCTTGTGTTGGGAAACGGACTCCGACAGTACTACTACGCCAAC  
TTGCGCAGGGGCTGCTCATGGAGCCACCGTCGACTGTGCCACCGGGTCTACTGGGACAATGGAGACTGCGCCGAC  
GCGGAGCCGGAGCTGATGTGCGCTCTGGAGTTAC **TAC**

**TaCBFIIIc-3.1 (EF028758)**

ATG GACATGGGCCTTGAGGTCTCGAGCTCTCCCGTCTCTCTCGTCCGGTGTCTCTCGCCCGAGCACGCGCGGG  
GCGGGCGTCTGCGCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCCGTGTACCGCGCGGTGC  
GCGCGCGGGCAACGCCAAGCTTGGGTCTGCGAGGTGCGCGTCCCGGCAAGCGCGCGCGAGGCTCTGGCTCGGG  
GCTACGACACGGCTGAGCTCGCAGCGCGCGGGAACGACGCGCGCATGCTTGCCTGGGCGGCGCTCCGCGCGGTG  
GTCAACTTTCGGGACTCCGCGTGTTGCTCGCGGTGCCGTCCGCACTCTCCGACCTCGCGGACGTCCGGCGCGCGG  
CGGTGAGGCGCGTGGCGAATTGACGACGAAAGGCTGGCAACGGCTCCCTACCGCCACCGTACCGAAGAGGCC  
TCTGTGACGCTCTGAAGAATCGTCTGAGTCTGACAGTGCCGCTTGTGCGGAGACGTGGAACCTTCAGCCGA  
TAGAGAGTTCGAGGTGCCGTCGAGTGGACACCGATATGTTCCGGCTTGACTTGTTCGGGAAATGGACCTGTGCT  
CGTACTACGCGAGCCTCGCGGAGGCACTGCTCGTGACCCGCCGCGACGGGTGACCACACCGACACGTACTGGGAC  
AACGGCGACGGCGGAGCCGATGTGCGCTCTGGAGTAC **TAC**

**TaCBFIIIc-3.2a (EF028759)**

ATG GACATGGGCCTTGAGGTCTCGAGCTCTCCCGTCTCTCTCGTCCGGTGTCTCTCGCCCGAGCACGCGCGGG  
GCGGGCGTCTGCGCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCCGTGTACCGCGCGGTGC  
GCGCGCGGGCAACGCCAAGCTTGGGTCTGCGAGGTGCGCGTCCCGGCAAGCGCGCGCGAGGCTCTGGCTCGGG  
GCTACGACACGGCTGAGCTCGCAGCGCGCGGGAACGACGCGCGCATGCTTGCCTGGGCGGCGCTCCGCGCGGTG  
GTCAACTTTCGGGACTCCGCGTGTTGCTCGCGGTGCCGTCCGCACTCTCCGACCTCGCGGACGTCCGGCGCGCGG  
CGGTGAGGCGCGTGGCGAATTGACGACGAAAGGCTGGCAACGGCTCCCTACCGCCACCGTACCGAAGAGGCC  
TCTGTGACGCTCTGAAGAATCGTCTGAGTCTGACAGTGCCGCTTGTGCGGAGACGTGGAACCTTCAGCCGA  
TAGAGAGTTCGAGGTGCCGTCGAGTGGACACCGATATGTTCCGGCTTGACTTGTTCGGGAAATGGACCTGTGCT  
CGTACTACGCGAGCCTCGCGGAGGCACTGCTCGTGACCCGCCGCGACGGGTGACCACACCGACACGTACTGGGAC  
AACGGCGACGGCGGAGCCGATGTGCGCTCTGGAGTAC **TAC**



CCTCAACTTCGCGGACTCCGCGTGGCTGCTCGCCGTGCCGCCGCACTCGCCGACCTCAGCGACGTCCGGCGCGG  
 GGCGGTTCGAGGCCGTGCTGATTCCAGAGACGAGAGGCTGCCAACGGCTCCCTCACAGTCACCGCCACCGTCACT  
 AAGAGGCCTCCTGTGGCGCTCCTGAAGAATCGTCTGAGTCTGACAGTGTGCGTTCGTCGGAGACGTGCGAACT  
 TCTGCCGATGGAGAGTTCGAGGTGCCGGTCCGCGGTGGACACCGATATGTTTCAAGGCTTGACTTGTTCCTCGGA  
 CCTGTGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCCTGGACCCGCCGGCACCGGTGACCACCACCTACGCGT  
 ACTGGGACAACGGCGACGGCGGAGCTGATGTGCGGCTCTGGAGCTACTAG

>TaCBFIIIc-3.2b (Accession)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCGTCTCCTCGTTCGGCGTCTCCTCGCCCCGAGCACGCGGCGGG  
 GCGGGCGTTCGTGGCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGTTGTACCGCGGCGTGC  
 TGTACCGGGGCAACGCGGAGCGGTGGGTCTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGGAGGCTCTGGCTCGGG  
 TGTACCGGCAACGCGGTGAGATCGCAGCGCGCGGCAACGAGCGCGGCGCATGCTCGCCCTGGGCGGCGGCTCCGCGCGGTG  
 CTCAACTTCGCGGACTCCGCGTGGCTGCTCGCCGTGCCGCCGCACTCGCCGACCTCGGCGACGTCCGGCGCGCGG  
 CGGTTCGAGGCCGTGCTGATTTCAGAGACGAGAGGCTGCCAACGGCTCCCTCACAGTCACCGCCACCGTCAACGAA  
 GAGGCCTCCTGTGGCGCTCCTGAAGAATCGTCTGAGTCTGACAGTGTGCGTTCGTCGGAGACGTGGAACCTTC  
 TCGCGATGGAGAGTTCGAGGTGCCGGTCCGCGGTGGACACCGATATGTTTCAAGGCTTGACTTGTTCCTCGGA  
 TGTGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCCTGGACCCGCCGGCACCGGTGACCACCACCTACGCGTAC  
 TGGGACAACGGCGACGGCGGAGCTGATGTGCGGCTCTGGAGCTACTAG

>TaCBFIIIc-D3 (EF028760)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCATCTCCTCGTTCGGTGTCTCCTCGCCCGTGCACGCGGCGGG  
 GCGGGCGTTCGTGGCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGTTGTACCGCGGAGTGC  
 TGTACCGGGGCAACGCGGAGCGGTGGGTCTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGGAGGCTCTGGCTCGGG  
 TGTACCGGCAACGCGGTGAGGTGCGCGGCGCGCGGCAACGAGCGCGGCGCATGCTTGCCTTGGGCGGCGGCTCCGCGCGGTG  
 CCTCAACTTCGCGGACTCCGCGTGGCTGCTCGCCGTGCCGCCGCACTCTCCGACCTCGGCGACGTCCGGCGCGCGG  
 CGGTTCGAGGCCGTGCGGATTTCCAGAGGCGAGAGGCTGCCAACGGCTCCCTCACCGCCACCGTCAACGAAGAGGCC  
 TCTGTGGCGCTCCTGAAGAATCGTCTGAGTCTGACAGTGC CGGTTTCGTCGGAGACGTGCGAACCTTCTGCCGA  
 TGCAGAGTTCGAGGTGCTGTCGAGTGGACACCGATATGTTTCAAGTGGCTTGACTTGTTCCTCGGAATGGACCTGT  
 GCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCGTGGACCCGCCGTGACGGTGGCCATCATGACTCGTACTGG  
 GACAACGGCGACGACGGAGCTGATGTGCGGCTCTGGAGCTACTAG

>TaCBFIIIc-B10 (EF028761)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCATCTCCTCGTTCGCTGGCCAAGCGCCCCGCGGGGCGCACCAA  
 GTTCCCGGAGACGCGGCACCCCTGTGTATCGTGGCGTGC CGGCGCGGGGCAACGCGCAACGGTGGGTCTGCGAGGTGC  
 TGTACCGGCAACGCGGTGAGGTGCGCGGCGCGCGGCAACGAGCGCGGCGCATGCTTGCCTTGGGCGGCGGCTCCGCGCGGTG  
 CCTCAACTTCGCGGACTCCGCGTGGCTGCTCGCCGTGCCGCCGCACTCTCCGACCTCGGCGACGTCCGGCGCGCGG  
 CGGTTCGAGGCCGTGCGGATTTCCAGAGGCGAGAGGCTGCCAACGGCTCCCTCACCGCCACCGTCAACGAAGAGGCC  
 TCTGTGGCGCTCCTGAAGAATCGTCTGAGTCTGACAGTGC CGGTTTCGTCGGAGACGTGCGAACCTTCTGCCGA  
 TGCAGAGTTCGAGGTGCTGTCGAGTGGACACCGATATGTTTCAAGTGGCTTGACTTGTTCCTCGGAATGGACCTGT  
 GCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCGTGGACCCGCCGTGACGGTGGCCATCATGACTCGTACTGG  
 GACAACGGCGACGACGGAGCTGATGTGCGGCTCTGGAGCTACTAG

>TaCBFIIIc-10.1 (Accession)

ATGGCATGGGCCTTGAGATCCTGAGCTCCTCCCCCTCCTCTTCCAACGAGAACGCGTTGGTGGCCAAGCGCCCCGGC  
 GGGGCGCACCAAGTTCCGGGAGACAAGGCACCCGTTGTACCGCGGCGTGC CGGCGCGGGGCAACGCGGAGCGGTGGG  
 TCTGCGAGGTGCGCGTCCCGGGCAAGCGCGGCGCGAGGCTCTGGCTCGGGACGTATGCCACGGCCGAGATCGCAGCG  
 TCGCGCAACGAGCGCGGCGCATGCTCGCCCTGGGCGGCGCGCTCCGCGCGCGCGCTTAACCTTCCCGGACTCCGCGTGGCT  
 GCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGCGGTGAGGCCGTGCGCGATTGTCAGCGACGGAGGCTG  
 GACGGGAGGCTGCCGTGGGTCCATCACCGCCACCGTCAACGAGGAGGCTCCTGTGGCGCTCTGCCGAGTCTGCTG  
 TCGGAGTCCGACGATCCCGGTTTCGTCGGAGACGTGAAACCTTCTGCGCATGGGGGCTTTGCGCTCGCGGCGGGAAT  
 GGACGTGAAATGTTTCAAGTGGCTTGACTTGTTCCTCGGAACGAGCTTGGGCTCGTACTACGCGAGCCTCGCGGAGG



CGCTGCTCATGGACCCACCGCCGGTGGCGACCGGCACCGGCGCTACTGGGATGACGGAGAGTTCCGGCAGGTGG  
CAACTGAGTTCGCGCTCTGGAGCTTCTAG

>TaCBFIIIc-10.2 (Accession)

ATGGGCATGGGCTTGAGATCTCGAGCTCCTTCCCTTCTCTTCCAACGAGAACGCGTTGGTGCCCAAGCGCCCGG  
GGGGCGCACCAAGTTCCGCGAGACGCGGCACCCGCTGTACCGCGCGTGGCGCGCGGGCAACGCCCAACGGTGGG  
TCTGCGAGGTGCGCGTCCCTGGCAAGCGCGCGCTCGGCTCTGGCTCGGGACGTACGCCACAGCCGAGATCGCAGCG  
CGCGCCAACGACGCGCGCATGCTCGCCCTGGCGCGCGCTCCGCCGCGCGTCTCAACTTCCCGGACTCCGCGTGGCT  
GCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGCGCGTCTGAGGCCGTCCGCGATTTCAGC  
GACGGGAGGCTGCCGTGGGTCCATCACCGCCACCGCCACCGCCGAGGAGACCTCCTGTGGCGCTCCTGCGGAATCC  
TCGTCTGAGTCTGACGATGCCGTTTCGTGAGAGACGTGAAACCTTCTGGCAATGGAGACTTCGCGCTGCCGGGCGG  
AATGGACGTGAAATGTTTCACTCGGCTGACTTGTTCGGGAAATGGACTTCGGCTCGTACTACGCGAGCCTCGCGG  
AGGCGCTGCTGATGGACCGCGCGGTGGCGACCGGCACCGGGCGTACTGGGATAACGGAGAGTTCGCGCGAGGGG  
GCAACTGATTTCGCCCTCTGGAGCTATAG

>TaCBFIIIc-13.1a (Accession)

ATGGATCTTTCCAGCTCCTTCCCGTCTCGTGGCGTCATCCTCGCCCGAGCACGCGTCGGGGCGGGCGTCGCCGGC  
CAAGCGCCCCCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGCTGTACCGCGGTGTGCGAGGTGCGCGTCCCT  
CGCAAGCGTGGTTCTCGGCTCTGGCTCGGGACGTACCTGACGGCCGAGGCCCGCGCGCGCGCAACGACGCGGCCAT  
CTCGCACTCGGCGGCGCTCCGCCAGGTGCCTCAACTTCGCGGACTCGGCGTGGCTGCTCGCCGTGCCGTCCGCGC  
TCTCCGACCTCGCCGACGTCCGGCGCGCGGCCCTGCAGGCCGTGCGGATTTCAGCGATGGGAGGCTGCCAATGGC  
CTGGTAACAGGACCGCGCGTGAAGAGGCCCATCTAGCGCTCCTGCACAATCGTCGTCTGAGTCCGACAGCGCCGA  
TTGCTCGGAGACGTGGAAGCTTCCGCCGATGGAGAGTTCGAGGTGCTGGCTACAATGGACATCGACATGTTTCAGGC  
TTGACCTGTTCCCGGAAATGGACCTGGGCTCGTACTACGTGAGCCTCGCGGAGGCGCTGCTCATGGACCGCGCGTGC  
ACGGCCACCATCATAGACGCGTACCGGGACAACCGCGACGGCGGTGCTGATGTCGCGCTCTGGAGCTACTAG

>TaCBFIIIc-13.1b (Accession)

ATGGATCTTTCCAGCTCCTTCCCGTCTCGTGGCGTCATCCTCGCCCGAGCACGCGTCGGGGCGGGCGTCGCCGGC  
CAAGCGCCCCCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGCTGTACCGCGGTGTGCGAGGTGCGCGTCCCT  
CGCAAGCGTGGTTCTCGGCTCTGGCTCGGGACGTACCTGACGGCCGAGGCCCGCGCGCGCGCAACGACGCGGCCAT  
CTCGCACTCGGCGGCGCTCCGCCAGGTGCCTCAACTTCGCGGACTCGGCGTGGCTGCTCGCCGTGCCGTCCGCGC  
TCTCCGACCTCGCCGACGTCCGGCGCGCGGCCCTGCAGGCCGTGCGGATTTCAGCGATGGGAGGCTGCCAATGGC  
CTGGTAACAGGACCGCGCGTGAAGAGGCCCATCTAGCGCTCCTGCACAATCGTCGTCTGAGTCCGACAGCGCCGA  
TTGCTCGGAGACGTGGAAGCTTCCGCCGATGGAGAGTTCGAGGTGCTGGCTACAATGGACATCGACATGTTTCAGGC  
TTGACCTGTTCCCGGAAATGGACCTGGGCTCGTACTACGTGAGCCTCGCGGAGGCGCTGCTCATGGACCGCGCGTGC  
ACGGCCACCATCATAGACGCGTACCGGGACAACCGCGACGGCGGTGCTGATGTCGCGCTCTGGAGCTACTAG

>TaCBFIIIc-13.1c (Accession)

ATGGATCTTTCCAGCTCCTTCCCGTCTCGTGGCGTCATCCTCGCCCGAGCACGCGTCGGGGCGGGCGTCGCCGGC  
AAGCGCCCCCGGAGAGCGCACCAAGTTCCGGGAGACGCGGCACCCGCTGTACCGCGGTGTGCGAGGTGCGCGTCCCT  
GGGGCGGTGGCTGTGCGAGGTGCGCGTCCCTGGCAAGCGTGGTTCTCGGCTCTGGCTCGGGACGTACCTGACGGCCG  
AGGCCGCGCGCGCGCAACGACGCGCGCATGCTCGCACTCGGCGGCGCTCCGCCAGGTCACTCAACTTCGCGGAC  
TCGGCGTGGCTGCTCGCCGTGCCGTCCGCGCTCTCCGACCTCGCCGACGTCCGGCGCGCGCGGCCCTGCAGGCCGTGCG  
GGATTTCAGCGATGGGAGGCTGCCAATGGCCTGGTAACAGGACCGCGCTGAGGAGGCCCATCTAGCGCTCCTG  
CACAATCGTCGTCTGAGTCCGACAGCGCGGATTCTGTCGGAGACGTGGAAGCTTCCGCCGATGGAGAGTTCGAGGTG  
CTGGCTACAATGGACATCGACATGTTTACGGCTTGACCTGTTCCCGGAAATGGACCTGGGCTCGTACTACGTGAGCCT  
CGCGGAGGCGCTGCTCATGGACCGCGCTCGACGGCCACCATCATAGACGCGTACCGGGACAACCGCGACGGCGGTG  
CTGATGTCGCGCTCTGGAGCTACTAG

>TaCBFIIId-12.1 (EF028762)

ATGGACACGGGCGCGGAGCGCAACTGGAACCTCGCCGGCATCCCCGCGTCTCTGCTCGAGCAGGGGATGCCGACGTC  
GCCTGCGTCGCCGACGCCAAGCGCCCCCGGGGCGAACCAAGTTCAAGGAGACGCGCCACCCGCTGTTCACGGCG



TCCGCGTCCGGGCGAGCAACGGCCGGTGGGTGTGCGAGGTGCCGCTGCCGGGAAGCGCGGCGAGCGGCTCTGGC  
 TGGCAGGCACGTACCGCGCGAGCGCGCGCGCGCGCACAGCGCGCCATGCTCGCGCTGTACGGCGCACCCCG  
 GCGCGCGGCTCAACTACCCCGACTCGGCGTGGCTGCTCGCCGTGCCCTCCTCCCTCTCCGACCTGGCGGACGTCCG  
 GCGCGCCGCATCGGGGCGGTGGTGGACTTCCTGCGCCGGCAGGAGGCGGGCGCCAGCGCCGGCGCGGTCTGCTGAGG  
 AGGCCACGTGACGGGATCGCCTCCGCCGCGTCCGGCGCAGACAATGCCAGCTCGTCCGGCGGCGGCGCCACTCA  
 CAGCCGCATGCGCCAATGCGGGGTACGAGGTGCCGGACGCATTGTGCCACGACATGTTGAGCTCCACAGCTCGGG  
 CGAAATGGACGCGGGCACGTACTACGCGACCTGGCGCAGGGGTGCTCCTGGAGCCGCGCGCGCGCTCCAGCG  
 GGGCGAGCTCGGAGCGGGAGACGATGCGGCGCTATGGAACCACTGA

>TaCBFIIId-B12 (EF028763)

ATGGACACGGGCCCCGAGCGCAACTGGAACCTCGCCGGCATCCCCGCGTCTCGCTCGAGCAGGGGATGCCGACGTC  
 CCTGCGTCCCGACGCGGAAGCGCCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCAGTGTTCACGGCG  
 TCGCGCGCGGGGTAGTAACGGCCGATGGGTGTGCGAGGTGCGCGTGCCGGGAAGCGCGCGCAACGGCTCTGGCTC  
 GGCACGCACGTACCGCGGAGGCGGCGCGCGCGCACGACGCGCGCATGCTCGCGCTGTACGGCGCACCCCGCG  
 TCGCGCGCTCAACTACCCCGACTCGGCGTGGCTGCTCGCCGTGCCCTCCTCCCTCTCCGACCTGGCGGACGTCCGGC  
 GCGCGCCCATCGGGGCGGTCTGTTGACTTCCTGCGCCGGCAGGAGGCGGGCGCTAGCACGGCGCGGTACCGAGGTG  
 GCCTCCATCGACGGGATCGCCTCCGCCGCGTCCGCACAGACAATGCCAGCTCGCGGGCGGCGGCATCCCACTCCCA  
 GCGGCCATGCGCCAATGCGGAGTTCGAGGTGCCGGACGCATTGTGCCACGACATGTTGAGCTCCACAGCTCCGGCG  
 AAATGGACGCGGGCACGTACTACGCGACCTCGCGCAGGGGTGCTCCTGGAGCCGCGCGCGCGCTCCAGCGGG  
 GCGAGCTCGGAGCAAGGAGATGATGCGGCGCTATGGAACCACTGA

>TaCBFIIId-A27 (JF758493) (MacLachlan, Baga and Chibbar 2011 genomic 5A)

ATGGACACGGGCCCCGAGCACAACTGCAGCTCGCCGGCTTCCTGCGCGTCTCGCAGGGGAGGTGATGCCGACGTC  
 GCCACGTGCGCGACGCGGAAGCGCCCCGCGGGGCGCACCAAGTCAAGGAAACAGGCCACCCGTTGTACACGGCG  
 TCGCGCGCGGGGCGGCAACGGCCGGTGGGTGTGCGAGATGCGGGTGCCCGGAAGCGCGGCGAGCGGCTCTGGCTC  
 GGCACGCACGTACCGCGCAAGGCGGCGCGCGCGCGCACGACGCGAGCCATGCTCGCGCTGCACGGCGGCTCCGGCGG  
 GCGCCTCAACTTCCTTACTCCGCGTGTCTTCTCGCCGTGCCCTCCTCTCTCTAGCCTGGCAGACGTCAGGCGCG  
 CCGCCATCGGGGCGGTCTGTTGACTTCCTGCGCCGGCAGGCCACGATCGCCGGCGCCAGGGGTGCCGAGGTAGTTCCC  
 GTCAACGGGGTTCGCTCCGTAGCGCCCGCGCGGGCAATGCCAGGTCTGCGCGACGTCTGCCAGCAGCCTTGTGC  
 CAATGCTGAGTCCGAGGCGCCTGACGCATTGCGCGGCGGCTGCCGGAGCTCCACACTTCCGGGAAATGGACGTGA  
 GCACGTACTACGCGACCTTGCAGGGGCTTCTCTGGAGCCGCGCCACCAGCGGCGGCGAGGACTGCAACGACGGC  
 GGAGATGATGCGGTGCTATGGAGCCACTGA

>TaCBFIIId-D27 (JF758490) (MacLachlan, Baga and Chibbar 2011 genomic 5D)

ATGCACACGGGCCCCGAGTACAAGTGGAGCTCGCCGGTTTCGTCGCGTCTCGCAGGAGCAGGGGATGCCGACGTC  
 GCCACGTGCGCGACGCGGAAGCGCCCCGCGGGGCGCACCAAGATCAAGGAGACAGGCCACCCGTTGTACACGGCG  
 TCGCGCGCGGGGCTGCAACGGCCGGTGGGTGTGCGAGATGCGGGTGCCCGGAAGCGCGGCGAGCGGCTCTGGCTC  
 GGCACGCACGCCACCGCGCAAGGCGGCGCGCGCGCGCACGACGCGAGCCATGCTCGCGTTGCACGGCGGCTCCGGCGG  
 GCGCCTCAACTTCCTTACTCCGCGTGTCTGCTCGCCGTGCCCTCTTCTCCCTCCAGCCTGGCCGACGTACAGGCGCG  
 CCGCCATCGGGGCGGTCTGTTGACTTCCTGCGCCGGCAGGCCAGGACCGCGGCGCGGCGACGCCGAGGTAGTTCCC  
 GTCAACGGGGTTCGCTCCGTAGCGCCCCCGCGGGCAATGCCAGGTCTGCGCGACGTCTGCCAGCAGCCTTGTGC  
 CAATGCTGAGTCCGAGGTGCTTACGCATTGCGCGGCGGCTGCCGGAGCTCCACAGTCCGGGAAATGGACGTGA  
 GCACGTACTACGCGACCTTGCAGGGGCTGCTCCTGGAGCCGCGCCACTGGCGGCCAGTACTGCAACGACGGC  
 GGAGATGATGCGGCGCTATGGAGCCACTGA

>TaCBFIIId-A15 (EF028764)

ATGGACATGACCGGCTCCGATCAGCAGCGGAGCTCCCCGTCTCGCCGTCTCCTCCTCGCACCTGAAGCGTCCCGC  
 CGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGTTGTACCGCGGCTGCGGCGCGCGGCGAGCGCGCGGCGGTGGG  
 TGTGCGAGGTGCGCGTCCCGCGCAAGCGCGGCGAGCGGCTGTGGCTCGGCACGCACCTCACTGCCGAGGCGGCGCGG  
 GAGCGTACGACGCGCGCATGCTCTGCCTGATCGGCGCCTCCACCCAGTGTCTCAACTTCGCGACTCCGCGCTGGCT  
 CCTCGCGGTGCCGTCCGCGCTCCCGACTTTCAGACGTCCGGCGCGCGGCTCTCAGCGTGTGCGAGACTTCCAGC  
 GCGGGGAGGTGCCAGCGGCGCGCCACCAGGTCCCTGGACGCCACCGTTCCCGTCGATGATGGAACCTGTAGCCAA



TCCGCGCAGTCGTCCATGGAAAATACCGGCTCATCATGGACGTCGTCTCCCTGCCGTCTGGCAACGGAATGTTT  
GAGGTGCCGGCCACACTAGGCTGCGACATGTTTCGAGCTGGATATGTCCGGGAAATGGACCTGGACACATACTATGC  
GTACTTCGCGGAGGGGCTTCTCCTCGAGCCGCCCAACCGCCGGTCGCCGGCGGTGCTGGACACCGAAGGCGGCG  
GAGCAGACGCCGCGCTCTGGAGCTAC**TAA**

-----15- (Accession)  
GACATGACCGGCTCCGACCAGCAGCGGAGCTCCCCGTCTCGCCGTCTCTCCTCGCAC**CTGAAGCGTCTTG**  
CGCGCGCACCAGTTCAAGGAGACCGCCACCCG**GTGTACCGCGCGTGGCGCGCGGGGCAACGCCGGCGGT**  
TGTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGAGCGGCTCTGGCTCGGCACGCACCTCACTGCCGAGCGCGCGCG  
CGAGCGGACGAC**CCCGCCATGCTCGGCTGATCGGCCCTCCACCCCGTGCTCAACTTCGTCGACTCCGCTTG**  
CTCGCCGTGCCGTCCGCGCTCTCCGACTTTGCAGACGTCCGGCGCGCGGCTCTCAGCGCTGTTCGCGGACTTCCAAC  
GCCGGGAGGCCACCAGCGCGCGGCCACCACGTCCCTGGCCGCCACCGTCCCCGTGATGACGGAAGCTGTAGCCAA  
TCCGCGCAGTCGTCCATGGAAAATACCGGCTCGTCTGGAGCTCGTCCCTACCTTCTGGCGACGGAATGTTTGGGT  
GTTTGGGTGCCGGCCACACTGGGCTGCAACATGTTTCGAGCTGGATATGTCCGGGAAATGGACCTGGACACATACT  
ATGCGTACTTTCGCGGAGGGGCTTCTCCTCGAGCCGCCCAACCGCCGGTCGCCGGCGGTGCTGGGACACCGAAGGC  
AGCGGAGCAGACGCCGCGCTCTCGAGCTAC**TAA**

>TaCBFIIId-15.2b (Accession)

ATG**GACATGACCGGCTCCGACCAGCAGCGGAGCTCCCCGTCTCGCCGTCTCTCCTCGCACCTGAAGCGTCTTG**  
CGCGCGT**TACCAAGTTCAAGGAGACCGCCACCCG**GTGTACCGCGCGTGGCGCGCGGGGCAACGCCGGCGGT  
TGTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGAGCGGCTCTGGCTCGGCACGCACCTCACTGCCGAGCGCGCGCG  
CGAGCGGACGAC**CCCGCCATGCTCGGCTGATCGGCCCTCCACCCCGTGCTCAACTTCGTCGACTCCGCTTG**  
CTCGCCGTGCCGTCCGCGCTCTCCGACTTTGCAGACGTCCGGCGCGCGGCTCTCAGCGCTGTTCGCGGACTTCCAAC  
GCCGGGAGGCCACCAGCGCGCGGCCACCACGTCCCTGGCCGCCACCGTCCCCGTGATGACGGAAGCTGTAGCCAA  
TCCGCGCAGTCGTCCATGGAAAATACCGGCTCGTCTGGAGCTCGTCCCTACCTTCTGGCGACGGAATGTTTGGGT  
GCCGGCCACACTGGGCTGCAACATGTTTCGAGCTGGATATGTCCGGGAAATGGACCTGGACACATACTATGCGTACT  
TCGCGGAGGGGCTTCTCCTCGAGCCGCCCAACCGCCGGTCGCCGGCGGTGCTGGGACACCGAAGGC**GGCGGAGCA**  
GACGCCGCGCTCTGGAGCTAC**TAA**

>TaCBFIIId-16.1a (Accession)

ATG**GACATGACCGGCTCGGACCAGCAATGGAGCTCCTCTTCTCGCCGTCTGTCGACCTCCTCGCACCGGAAGCGCCC**  
CGCCGGGCGCACCAAGTTCAAGGAGACCGCCACCCG**GTGTACCGCGCGTGGCGCGCGGGGCAACGCCGGCGGT**  
GGTGTGCGAGGTGCGCGTCCCCGGCAGCGCGCGGAGCGGCTCTGGCTCGGAACGTACCTCACAGCGGACCGCGCG  
CGCGGTGCGCACGACCGCGCCATGCTCGGCTGCTCGGTGCTCGCGCGGTGCTCAATTT**CGCGGACTCCGCTTG**  
GCTCCTTGCCGTGCCGCGCGCTCGCCGACCTCGCGGCCGTGAGCGCGCGGCCCTCGCCGCCGTGGCGGACTTCC  
AGCGCCGGCATGCCCTCCAACAGCGCAGCCACCGTCCCCGTGATGAGGAGACCTCCGGCGCGTCCGCTCTGTCTGTCT  
GCGGACAATGCGAGCGGTTCTGTAGCGACGTGCAACCTTGGGCTGAGGGAACGTTCGAGGTGCCGTCCGCGCTGGG  
CAGCGACATGTTTCGAGCTGGACTTGTCCGGGAGATGGACCTGGGCACGTACTACGCGGACCTCGCGGACGGGCTGC  
TCCTGGAGCGCGCGCGTCTGTGGACAGCGGGGCGTGTGGGACACCGGAGACGGCGGAGCTGACTCCGGGCTCTGG  
AGCTAC**TGA**

>TaCBFIIId-16.1b (Accession)

ATG**GACATGACCGGCTCGGACCAGCAATGGAGCTCCTCTTCTCGCCGTCTGTCGACCTCCTCGCACCGGAAGCGCCC**  
CGCCGGGCGCACCAAGTTCAAGGAGACCGCCACCCG**GTGTACCGCGCGTGGCGCGCGGGGCAACGCCGGCGGT**  
GGTGTGCGAGGTGCGCGTCCCCGGCAGCGCGCGGAGCGGCTCTGGCTCGGAACGTACCTCACAGCGGACCGCGCG  
CGCGGTGCGCACGACCGCGCCATGCTCGGCTGCTCGGTGCTCGCGCGGTGCTCAATTT**CGCGGACTCCGCTTG**  
GCTCCTTGCCGTGCCGCGCGCTCGCCGACCTCGCGGCCGTGAGCGCGCGGCCCTCGCCGCCGTGGCGGACTTCC  
AGCGCCGGCATGCCCTCCAACAGCGCAGCCACCGTCCCCGTGATGAGGAGACCTCCGGCGCGTCCGCTCTGTCTGTCT  
GCGGACAATGCGAGCGGTTCTGTAGCGACGTGCAACCTTGGGCGGAGGGAACGTTCGAGGTGCCGTCCGCGCTGGG  
CAGCGACATGTTTCGAGCTGGACTTGTCCGGGAGATGGACCTGGGCACGTACTACGCGGACCTCGCGGACGGGCTGC  
TCCTGGAGCGCGCGCGTCTGTGGACAGCGGGGCGTGTGGGACACCGGAGACGGCGGAGCTGACTCCGGGCTCTGG  
AGCTAC**TGA**





ATG GACTTTGGTATCAACGGCTGGATCAGCTCCCCTTCCTCATCGACGTCCGGGCACGAGCTCGGGGATGCGGTG  
 CCCGTGTGGTTGCCGGCG **GCGAAGCGGCCCGCAGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG** GTGTACCGTGG  
 CTGCGGCGCGGGGCGAGCGCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGC  
 TCGGACGTACGTGCGCGCGAGTCCGCGCGCGCTCTCAGACGCGCCATGCTCGCCCTGCTCGGACGCTCCCCC  
 TCGCGCGCGGCGTGCCTCAACTTCCCG **GACTCCGCGTGGCTG** CTGGTCATGCCCCGAGGCTCTCCGACCTGGCCGA  
 CGTCCGGCGCGCGCCATCCAGGCCGTGCGGGCTTCTGCGCTGGAGGCCGCCACCGTGGTCCCGACGTGACG  
 AGGCCACCTCCCCTGTGTACCTGCCGTGCGCGTGGACAATGCGGACGAAGTGTTCAGGTGCCGACTTTCTCCCCG  
 CTGGGACGACATGTTGAGCTCGACATGTCCGGGAAATGACCTGGACGCGTACTACGCGGGCTTTGCCAGGG  
 GATGCTCCTGGAGCCGCCACGCCAGCGTACTGGGAGAACGGAGAGTGCGGCGACGGCGGAGCGGCCGCCGCC  
 TCTGGAGCTAC **TGA**

>TaCBFIIId-B19 (EF028767)

ATG GACATGGGCATCAACGGCTGGATCAGCTCCCCTTCCTCATCGACGTCCGGGCACGAGCTCGGGGAGACCGTGCC  
 TGTGTGGTCGCCGGCG **GCGAAGCGGCCCGCAGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG** GTGTACCGTGGC  
 TCGCGCGCGGGGCGAGCGCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGCTC  
 GGGACGTACGTGCGCGCGAGTCCGCGCGCGCGCTCAGCAGCGCGCTATGCTCGCCCTGCTCGGACGCTCCCCCTG  
 TCGCGCGCGGTGCCTCAACTTCCCG **GACTCCGCGTGGCTG** CTGTCATGCCCCGAGGCTCTCCGACCTGGCCGACG  
 TCCGGCGCACGGCCATCCAGGCCGTGCGAGCTTCTGCGCTTGGAGGCCGCCACCGTGGTTCGGACGTGACGAG  
 GCCACCTCCCCCGTGTACCTGCCGTGCTCGTGGACAATGCCGACGAAGTGTTCAGGTGCCGACTTTCTCCCCGCT  
 GGGCAGCGACATGTTGAGCTCGACATGTCCGGGAAATGACCTGGACGCGTACTACGCGGGCTTTGCCAGGGGA  
 TGCTCCTGGAGCCGCCCGCCATGCCGTGCTACTGGGAGAACGGAGAAATGCGGCGACGGCGGCGCGGCCGCCGCC  
 TGGAGCTAC **TGA**

>TaCBFIIId-D19 (EF028768)

ATG GACATGGGCATCAACGGCTGGATCAGCTCCCCTTCCTCATCGACGTCCGGGCACGAGCTCGGGGAGGCGGTGCC  
 TGTGTGGTCGCCGGCG **GCGAAGCGGCCCGCAGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG** GTGTACCGTGGC  
 TCGCGCGCGGGGCGAGCGCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGCTC  
 GGGACGTACGTGCGCGCGAGTCCGCGCGCGCGCTCAGCAGCGCGCATGCTCACCCTGCTCGGACGCTCCCCCTC  
 TCGCGCGCGGTGCCTCAACTTCCCA **GACTCCGCGTGGCTG** CTGTCATGCCCCGAGGCTCTCCGACCTGGCCGACG  
 TCCGGCGCGCGGCCATCCAGGCCGTGCTGGGCTTCTGCGCTTGGAGGCCGCCACCGTGGTTCGGACGTGACGAG  
 GCCACCTCCCCCGTGTACCTGCCGTGCGCGTGGACAATGCCGACCAAGTGTTCAGGTGCCAATTTCTCCCCGCT  
 GGGCAGCGACATGTTGAGCTCGACATGTCTGGGAAATGACCTGGACGCGTACTACGCGGGCTTTGCCAGGGGA  
 TGCTCCTGGAGCCGCCCGCCACGCCGTGCTACTGGGAGAACGGAGAAATGCGGCGACGGCGGAGCGGCCGCCGCC  
 TGGAGCTAC **TGA**

>TaCBFIIId-24.1 (Accession)

ATG GACACGGCCATCGACAGCTGGATCAGCTCCCCTTCCTCATCGACGTCCGGGCACGAGCACGGGGAGGTGGTGCC  
 TGTGTGGTCGCCGGCG **GCGAAGCGGCCCGCGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG** GTGTACCGCGGCG  
 TCGCGCGCGGGGCGAGCGCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGCTC  
 GGGACGTACGCGCGCGCGAGTCCGCGCGCGCGCGCACGACGCGCCATGCTCGCGCTGCTCGGCCACTCCGCTC  
 TCGCGCGCGGTGCCTCAACTTCCCG **GACTCCGCGTGGCTG** CTGTCATGCCCCGAGGCTCTCCGACCTGGCCGATA  
 TCCGACGCGCGGCCATCGAGGCCGTGCGGTTTTCTGTGCTTGGAGGCCGCCGCGTGGTGCCCGTCTGTCGACGAG  
 GCCACCTCCCCCGTGTACCTGCTGTGCGCGTGGACAATGCCGACGAAGTGTTCAGTTGCCGACTTTCTCTGCCCA  
 GGGCAGCGACATGTTGAGCTCGACATGTCCGGGAAATGACCTGGACCCGTAACGCGGGCTTTGCCAGGGGA  
 TGCTCCTGGAGCCGCCCGCCACGCCGCGTACTGGGAGAACGGAGACGGCGGAGCGGCAGCCGGCTCTGGAGCTAC  
**TGA**

>TaCBFIIId-24.2a (Accession)

ATG GACATGGCCATCGACAGCTGGATCAGTTCCCCTTCCTCATCGACGTCCGGGCACGAGCACGGGGAGGTGGTGCT  
 TGTGTGGTCGCCGGCG **GCGAAGCGGCCCGCGGGCGCACCAAGTTCAAGGACACGCGCCACCCG** GTGTACCGCGGCG  
 TCGCGCGCGGGGCGAGTCCGCGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGCTC  
 GGGACGTACGTGCGCGCGAGTCCGCGCGCGCGCACGACGCGCCATGCTCGCGCTGCTCGGCCACTCCGCTC  
**TGA**



CGCCGCGGCGTGCTCAACTTCCCGGACTCCGCGTGGCTGCTCGTCATGCCCCCGTGGCTCTCCGACTTGGCTGA  
TATCCGACGCGCGGCCATCGAGGCCGTGCGGATTTTCTGTGCTGGAGGCCGCCCGTGGTGGCCATCATCGACG  
AGGCCACCTCGCCCGTGTACCTGCCGTGCCCGTGGACAATGCCTACGAAGTGTTCAGGTGCCGACTTCTCTGCG  
CAGAGCAGCGACATGTTTCGAGCTCGACATGTCCGGGAAATGGACCTGGACGCGTACTACGCAGGCTTTGCGCAGGG  
GATGCTCCTCGAGCCGCCGCCACGCCGACGTACTGGGAGAACGGAGAATGCGGCGACGGCGGAGCGGCCGCCGCC  
TCTGGAGCTACTGA

>TaCBFIIIId-24.2b (Accession)

ATGGACATGGCCATCGACAGCTGGATCAGCTCCCTTCTCATCGACGTCCGGGCACGAGCACGGGGAGGTGGTGTCT  
TGTGTGGTCCCGGCGCGAAGCGGCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGGCG  
TGGCGCGCGGGGCGAGTGGCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCCAAGCGCGGCGAGCGGCTCTGGCTC  
GGGACGTACGTGCGCGCGAGTGGCGCACGCGCGCGCACGACGCCACCATGCTCGCGCTGCTCGGCCACTCCGCCTC  
CGCCGCGGCGTGCTCAACTTCCCGGACTCCGCGTGGCTGCTCGTCATGCCCCCGTGGCTCTCCGACTTGGCCGATA  
TCCGACGCGCGGCCATCGAGGCCGTGCGGATTTTCTGTGCTGGAGGCCGCCCGTGGTGGCCATCATCGACGAG  
GCCACCTCGCCCGTGTACCTGCCGTGCGCCGTGGACAATGCCGACGAAGTGTTCAGGTGCCGACTTTCTCTGCGCA  
GAGCAGCGACATGTTTCGAGCTCGACATGTCCGGGAAATGGACCTGGACGCGTACTACGCAGGCTTTGCGCAGGGGA  
TGCTCCTCGAGCCGCCGCCACGCCGACGTACTGGGAAACGGAGAATGCGGCGACGGCGGAGCGGCCGCCGCCCTC  
TGGAGCTACTGA

>TaCBFIIIId-24.3 (Accession)

ATGGACACGGCCATCGACAGCTCGACAGCTCCCTTCTCATCGACTTCCGGGCACGAGCACGGGGAGGTGGTGTCT  
TGTGTGGTCCCGGCGCGAAGCGGCCCGCGGGGCGCACCAAGTTCAAGGAAACGCGCCACTCGCTGTACCGCGGCG  
TGGCTGCGCCGGGCGAGCGCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCCAAGCGTGGCGAACCGCTCTGGCTC  
GGGACGTACGTGCGCGCGAGTGGCGCGCGCGCGCACGACGCCCATGCTCGCGCTGCTCGGCCGCTCCGCCTC  
CGCCGCGGCGTGCTCAACTTCCCGGATTCCGCGTGGCTGCTCGTCATGCCCCCGTGGCTCTCCGACTTGGCCGATA  
TCCGACGCGCGGCCATCGAGGCCGTGCGGGTTTCTGCGCTGGAGGCCGCCCGTGGTGGCCATCGTCGACGAG  
GCCACCTCCCGCGTGTACCTGCCGTGCGCCGTGGACAGTCCGACAAAGTGTTCAGTTGCCGACTTTGTCTGCCCA  
GGGACGCGACATGTTTCGAGCTCGACATGTCCGGCGAAATGGACCTGGACGCGTACTACGCAGGCTTCGCGCAGGGGA  
TGCTCCTGGAGCCGCCGCCACGCCGCGTACTGGAAGAACGGAGAATGCGGCGACGGCGGAGCGGCAGCCGGCCTC  
GGGAGCTACTAA

>TaCBFIVa-A2 (EF028769)

ATGGACACCAACGCGCGCTGGCCGAGTTTGACGGGAAGAGTACAGGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTGTGGTCCGAGCCGCGAAGCGGCGGGCGGGGCGGAACAAGTTGACAGGAGACAGCCACCCAATGTACCGCG  
CGGTGCGCGCGCGTGGCGGGGAAGGCGAGTGGGTGTGCGAGCTGCGCGTGGCGGCCGGAAGCGCGAGTACTCCAGG  
ATCTGGCTTGGCACCTTCCGCGAGTGGCCAGATGGCGGCGCGCGCGCACGACTCGGCCCGCTCGCGCTCTCCGGCCG  
CGACGCGTGCTCAACTTCCCGGACTCCGCGTGGCGGATGATGCCCGTCCACGAGCCGGGTGCTTCAAGTTGGCCG  
CCGCGCAGGAGATCAAGGACGCCGTGCGCGTGGCCCTCAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGACGAGTCA  
ACGGCGCCGTGCTCCACGGCCGAGGAGAGCGCGCTCTCCATCATCCCAGCGACCTGTGCGGGCTCGACAATGAGCA  
CTGGATCGGCGGCATGGAGGCCGGGTGCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCCGCGCCGACG  
GAGCTTGGCGGGAGGACCGCGAACACGACGACGATTTCGACACGTGCTGTGGAGCTACTAG

>TaCBFIVa-2.2 (EF028770)

ATGGACACGATCGCCGCTGGCCGAGTTTGACGGGAAGAGTACAGGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTGTGGTCCGAGCCGCGAAGCGGCGGGCGGGGCGGAACAAGTTGACAGGAGACAGCCACCCAATGTACCGCG  
CGGTGCGCGCGCGTGGCGGGGAAGGCGAGTGGGTGTGCGAGCTGCGCGTGGCGGCCGGAAGTGGAGTTACTCCCGG  
ATCTGGCTTGGCACCTTCCGCGGTGCGCCAGATGGCGGCGCGCGCGCACGACTCGGCCCGCTCGCGCTCTCCGGCCG  
GGGCGCGTGCTCAACTTCCCGGACTCCGCGTGGAGGATGATGCCCGTCCACGCGGCCGGATGCTTCAAGTTGGCGG  
CCGCGCAGGAGATCAAGGACGCCGTGCGCGTGGCCCTCAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGACGTGTCA  
ACGGCTCCGTGCTCCACGGCCGAGGAGAGCGCGCTCTCCATCATCCCAGCGACCTGTGCGGGCTTGACAACGAGCA  
CTGGATCGGCGGCATGGAGGCCGGGTGCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCCGCGCCGACG  
GAGCTTCTTGGCGGGAGGACCGCGAACACGACGACGAGTTCGACACGTGCTGTGGAGCTACTAG



ATTAAGCAATGTAAAGTTCTAGAGAGTACTGCTAGTGCTAGATTGTGTTTCACCAAATATGGGAAGAACAGAGAG  
TAA

>TaCBFIVa-2.3 (EF028771)

ATG GACACCATCGCCGCTGGCCGAGTTTGACGGGCAAGAGTACAGGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTGTGGTTCGAGCCG CCGAAGCGCGGGCGGGCGGAACAAGTTGCAGGAGACACGCCACCCG GTGTACCGCG  
GCGTGGCGCCGCGTGGCCGGGAAGGGCAGTGGGTGTGGGAGCTGCGCGTGGCCGGCGGAAGCCGGAGTTACTCCAGG  
ATCTGGCTCGGCACCTTCGCCAGTGGCCAGATGGCGGGCGCGCGCACGACTCGGCCGGCGCTCGCGCTCTCCGGCCG  
CGACGCGTGCCTCAACTTCGCC GACTCCGCTGGCGGATGATGCCCCGTCCACGCGCCGGGTCTTCAAGTTGGCCG  
CCGCGCAGGAGATCAAGGACGCCGTGCGCGTGCCTCAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGACGAGTCA  
ACGGCGCCGTGCTCCACGGCCGAGGAGAGCGCGCTCTCCATCATCCCCAGCGACCTGTGGGGCTCGACAACGAGCA  
CTGGATCGGCGGCATGGAAGCCGGCTCGTACTACGCGAGCTTGCGCAGGGGATGCTCATGGAGCCCGGCCGACG  
GAGCTTCTTGGCGGGAGGACCGGAACACGACGACGGCTTCGACACGTGGCTGTGGAGCTAC TAG

>TaCBFIVa-A28 (JF758493) (MacLachlan, Baga and Chibbar 2011 genomic 5A)

ATG GACACCGTTGCCGCTGGCAGCAGTTTGAGGGGCAAGAGTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTGTGGTTCGAGCCG CCGAAGCGCGGGCGGGCGGATCAAGTTGCAGGAGACCGGCCACCCG GTGTACCGCG  
GCGTGGCGCCGCGTGGCAGGGGAAGGGCAGTGGGTGTGGGAGCTGCGCGTCCCCGTAAGCCGGGGTTACTCCAGGCTC  
TGGCTCGGCACCTTCGCCACCGCCGAGATGGCGGGCTCGCGCGCACGATTCCGCGCGCTCGCGCTCTCCGGCCACGA  
TGCCTGCCTCAACTTCGCC GACTCCGCTGGCGGATGATGCCCCGTCCACGCGACCGGGTCTTCAAGTTGCTCCCG  
CGAAGAGATCAAGGACGCCGTGCGCGTGCCTTCGAGGCATTCCAGGAGCAGCACCATGCCGACGCGTCCACGACG  
GAGGCGACGCGCCCTCCATCACCTCAAGTGACTTATCGGGCTGGACGACGAGCTCTTGATTGATGGCATGGACGC  
GGGGTCGTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCACCGGCCCGGTGCATGGCGGGAGGACCACG  
AACACGACGACGGCTTCGACACGCCGACGTCGCTGTGGAGCTAC TAG

>TaCBFIVa-D28 (JF758490) (MacLachlan, Baga and Chibbar 2011 genomic 5D)

ATG GACACCGTTGCCGCTGGCCGAGTTTGAGGGGCAAGAGTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTGTGGTTCGAGCCG CCGAAGCGCGGGCGGGCGGATCAAGTTGCAGGAGACCGGCCACCCG GTGTACCGCG  
GCGTGGCGCCGCGTGGCAGGGTTCGGCAGTGGGTGTGGGAGCTGCGCGTCCCCGTAAGCCGGGGTTACTCCAGACTC  
TGGCTCGGCACCTTCGCCACCGCCGAGATGGCGGGCTCGCGCGCACGACTCCGCGCGCTCGCGCTCTCCGGCCACGA  
TGCCTGCCTCAACTTCGCC GACTCCGCTGGCGGATGATGCCCCGTCCACGCGACCGGGTCTTCAAGTTGCCCCCG  
CGAAGAGATCAAGGACGCCGTGCGCGTGCCTTCGAGGCATTCCAGGAGCAGCACCCTGCAGACGCGTCCACGACC  
GAGGCGAGCGCCCTCCATCACCTCAAGTGACTTATCGGGCTGGACGACGAGCTCTTGATTGACGCGATGGACGC  
GGGGTCGTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGGCCCGGTGCATGGCGGGAGGACCAGC  
AACACGACGACGGCTTCGACACGCCGACGTCGCTGTGGAGCTAC TAG

>TaCBFIVa-25.1 (Accession)

ATG GACACCGTGGCGCCTGGCCGATTTTGAGGGGCAAGAGTACAGCACGGTGTGGCCGGAGGAGTACCGGAC  
GGTGTGGTTCGAGCCG CCGAAGCGCGGGCGGGCGGAACAAGTTGCAGGAGACCGGCCACCCG GTGTACCGCGCG  
TTCGCCGCGTGGCCGGGAAGGGCAGTGGGTGTGGGTGTGGGAGCTGCGGGTGGCCGGCGCCGGAGCCGGGTTTAC  
TCCAGGATCTGGCTCGGCACCTTCGCCGACCCGAGATGGCGGGCGCGCGCATGACTCGGCCGCGCTCGCGCTCTC  
CGGCCGCGATGCGTGCCTCAACTTCGCC GACTCCGCTGGCGGATGATGCCCCGTCCACGCGCCGGGTCTTCAAGC  
TGGCCGCCGCGCAGGAGATCAAGGACGCCGTGGCCGTGCGCTCAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGAC  
GTGTCAAAGGCGCCATCTCCACGGACTCCACGTCCGAGGAGAGCGCGCCGTCCATCACCTCCAGCGACTTGTCTGG  
TCTGGACGACGAGCACTGGATCGGCGGCATGGACGCCGGGTCTGACTACGCAAACCTTGGCGCAGGGCATGCTCATGG  
AGCCGCCGGCCCGGAGCGTGGCGGGAGGACCGGAACAGGACGACGGCGTCGACACGTGCTGTGGAGCTAC TAG

>TaCBFIVa-25.2 (Accession)

ATG GACACCGCGCGCCTGGCCGATTTTGAAGGGCAAGAGTACAGGACGGTGTGGCCGGAGGAGGAATACCGGAC  
GGTGTGGTTCGAGCCG CCGAAGCGCGGGCGGGCGGAACAAGTTGCAGGAGACCGGCCACCCG GTGTACCGCGCG  
TTCGCCGCGTGGCCGGGAAGGGCAGTGGGTGTGGGTGTGGGAGCTGCGGGTGGCCGGCGCCGGAGCCGGGTTTAC  
TCCAGGATCTGGCTCGGCACCTTCGCCGACCCGAGATGGCGGGCGCGCGCATGACTCGGCCGCGCTCGCGCTCTC  
TCCAGGATCTGGCTCGGCACCTTCGCCGACCCGAGATGGCGGGCGCGCGCACGACTCGGCCGCGCTCGCGCTCTC



CGGCCGCGATGCGTGCCTCAACTTCGCCGACTCCGCTGGCGGATGATGCCCCGTCCACGCGGCCGGGTTCGTTCAA  
 GTTGGCTGCCGCTCAGGAGATCAAGGATGCCGTGCGCTCGCCCTCAAGGCGTTCAGGAGCAGCAGCGCCCTGCCG  
 ACGCGTCAAAGCGCCATCCTCCACGGACTCGACGTCCGAGGAGAGCGCGCCGTCCATCACCTCCAGCGACTTGTCT  
 GGGCTGGACGACGAGCACTGGATCGCGGCGATGGACGCCGGGTCTACTACGCAAACCTTGGCGCAGGGCATGCTCAT  
 GGAGCCGCCGGCCCGGAGCGTGGCGGGAGGACCGCAACAGGACGACGGCGTCGACACGTCTGCTGTGGAGCTAC  
 AG

> TaCBFIVa-25.3 (Accession)

ATG GACACCGCCGGCGCTGGCCGCATTTTGAAGGGCAAGAGTACAGGACGGTGTGGCCGGAGGAGTACCGGAC  
 GGTGTGGTCCGAGCCGCGAAGCGCGGGCGGGCGGAACAAGTTGCAGGAGACGCGCCACCCG GTGTACCGCGGGC  
 TCGCGCCCGCTGGCCGGGAAGGGCAGTGGGTGTGGGTGTGCGAGCTGCGCGTGCCGGCCGCGGAAGCCGGGTTTAC  
 TCCAGGATCTGGCTCGGCACCTTCGCCGACCCCGAGATGGCGGCGCGCGCACGACTCGGCCGCGCTCGCGCTCTC  
 CGGCCGGGATGCGTGCCTCAACTTCGCCGACTCCGCTGGCGGATGATGCCCCGTGCACGCGGCCGGGTCTGTTCAAGT  
 TGGCTGCCGCTCAGGAGATCAAGGATGCCGTGGCCGTGCGCCCTGAAGGCGTTCCAGGGGACGAGCGCCCTGCCGAC  
 CGTTCGAAGGCGCCATCCTCCACGGACTCCACGTCCGAGGAGAGCGCGCCGTCCATCACCTCCAGCGACTTGTCTGG  
 GTTGGACGACGAGCACTGGATCGCGGCGATGGACGCCGGGTCTACTACGCAAACCTTGGCGCAGGGCATGCTTATGG  
 AGCCGCCGGCCCGGAGCGTGGAGGGAGGACCGCAACAGGACGACGGCGTCGACACATCGCTGTGGAGCTAC TAG

> TaCBFIVb-A20a (EF028772)

ATG GACACCGCCGCCCGCGCTCCCGCGTGAGGGGCACAGGACGGTGTGCTCCGAGCCG CCAAGAGGCCGGCAGG  
 CGCGACCAAGTTCAAAGAGACGCGCCACCCG CTGTACCGCGCGGTGCGCGCGCGGGCCGGCTCGGGCAGTGGGTGT  
 GCGAGGTGCGCGTGCGCGGCGCGCAAGGGTACAGGCTCTGGCTCGGCACCTTCACCACGGCCGAGATGGCGGCGCGC  
 GCGCACGACTCCGCCGTCTCGCGCTCCTCGACCCGCGCGCTGCCTCAACTTCGCCGACTCCGCTGGCGGATGCT  
 GCGCGTCTCGCGGCTGGCTCGTCCGCTTACGACGCGCGCGGAGATCAAGGACGCCGTGCGCGTCCGCGTTATGG  
 AGTTCCAGCGGACGCGCCCGTCTGTCCACGCCGAGACGCACGACGGCGAAAAGGACGTCCAAGGCTCACCAGCG  
 CCGAGCGAGCTGTCCACGTCCAGCGACTTGTGGACGAGCACTGGTTTGGCGGCATGAACGCCGGCTCGTACTATGC  
 GAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGCGCCAGAGCGCGGAGCGAGGATGGCGGCGAATACAGCGGCG  
 TCCAGACGCCGCTTTGGAACACTTATCCGACTAAT TAA

> TaCBFIVb-A20b (Accession)

ATG GACACCGCCGCCCGCGCTCCCGCGTGAGGGGCACAGGACGGTGTGCTCCGAGCCG CCAAGAGGCTGGCAGG  
 CGCGACCAAGTTCAAAGGAGACGCGCCACCCG CTGTACCGCGCGGTGCGCGCGCGGGCCGGCTCGGGCAGTGGGTGT  
 GCGAGGTGCGCGTGCGCGGCGCGCAAGGGTACAGGCTCTGGCTCGGCACCTTCACCAC TACCGAGATGGCGGCGCGC  
 GCGCACGACTCCGCCGTCTCGCGCTCCTCGACCCGCGCGCTTGCTCAACTTCGCCGACTCCGCTGGCGGATGCT  
 GCGCGTCTCGCGGCTGGCTCGTCCGCTTACGACGCGCGCGGAGATCAAGGACGCCGTGCGCGTCTGAGATTCC  
 AGCGGCAGCGGCCCGTCTGTCCACGCCGAGACGCACGACGGCGAAAAGGACGTCCAAGGCTCACCAGCGCCGCGC  
 GAGCTGTCCACGTCCAGCGACTTGTGGACGAGCACTGGTTTGGCGGCATGAACGCCGGCTCGTACTATGCGAGCTT  
 GGCGCAGGGGATGCTCATGGAGCCGCCGCGCCAGAGCGCGGAGCGAGGATGGCGGCGAATACAGCGGCGTCCAGA  
 CGCCGCTTTGGAACACTTATCCGACTAAT TAA

> TaCBFIVb-B20 (EF028773)

ATG GACACCGCCGCCCGCGCTCCCGCGTGAGGGGCACAGGACGGTGTGCTCCGAGCCG CCAAGAGGCTGGCAGG  
 CGCGACCAAGTTTAAGGAGACGCGCCACCCG CTGTACCGCGCGGTGCGCGCGCGGGCCGGCTCGGGCAGTGGGTGT  
 GCGAGGTGCGCGTGCGCGGCGCGCAAGGGTACAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGC  
 GCTCACGACTCCGCCGTCTCGCGCTCCTCGACCCGCGCGCTTGCTCAACTTCGCCGACTCCGCTGGCGGATGCT  
 GCGCGTCTCGCGGCTGGCTCGTCCGCTTACGACGCGCGCGGAGATCAAGGACGCCGTGCCATCGCCATCGTTGG  
 AGTTCCAGCGGACGCGCCCGTCTGTCCACGTCCGAGACGCACGACGGCGAGAAGGACGTCCAAGGCTCGCCGACG  
 CCGAGCGAGCTGTCCATGTCCAGCGACTTGTGGACGAGCACTGGTTTGGCGGCATGGACGCCGGCTCGTACTACGC  
 GAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGCGCCAGAGCGCGGAGCGAGGATGGCGGCGAATACAGCGGCGT  
 TCTACACGCCGCTTTGGAAC TAA

> TaCBFIVb-D20 (EF028774)







ATG GACGCCGCCGACGCCGCTCCCCGTGTGATGGCCACAGGACGGTGTGGTTCGGAGCCG CCGAAGCGGCCGGCC  
 GGCCGGACCAAGTTCAAGGAGACGCCACCCG CTGTACCGCGCGTGC GGCGACGGGGCCCCCGCGCCGGTGGGT  
 GTGCGAGGTGCGCGTGTCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCTGAGATGGCAGCGC  
 GCGCGCACGACGCCGCCGTTCTCGCGCTCTCTGGGCGCGCCGCTTGTCTCAACTTCGCC GACTCCGCCTGGCGGATG  
 CTCCCCGTCTCGCCGGCCCGTTTACGACCCGCCAAGGAGATCAAGGACGCCGTGCGCGTCTCGCGTCTTGGCGTTCCA  
 AAGGACGACCCGGTTCGCGTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCT  
 CGCCGGCGCCGAGCGCCCTGTTTCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCG  
 TTCTACTCGGAGGGCTGTTTCATGGAGTCGCCGGACACCAGACCGTGGCGGGAAGACCTCGAGCTCTGTGGCGTCGA  
 GACACCGCCATGGAGCTACTTGTTCGAC TAA

**TaCBFIVc-14.3 (EF028779)**

ATG GACGCCGTGACGCCGCTCCCCGTGTGATGGCCACAGGACGGTGTGGTTCGGAGCCG CCGAAGCGGCCGGCCGG  
 CCGGACCAAGTTCAAGGAGACGCCACCCG CTGTACCGCGCGTGC GGCGACGGGGCCCCCGCGCCGGTGGGTGT  
 GCGAGGTGCGCGTGTCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGC  
 GCGCACGACGCCGCCGTTCTCGCGCTCTCGGGCGCGCCGCTTGTCTCAACTTCGCC GACTCCGCCTGGCGGATGCT  
 CCCCCGTCTCGCCGGCCCGTTTACGACCCGCCAAGGAGATCAAGGACGCCGTGCGCGTCTCGCGTCTTGGCGTTCCAAA  
 GGCAGCACCCGGTTCGCGTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCTCG  
 GCGCGCCGAGCGCCCTGTTTCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGTGCGCATCGATGCCGGATCGTG  
 TACTCGGAGGGCTGTTTCATGGAGTCGCCGGACACCAGACCGTGGCGGGAAGACCTCGAGCTCGGTGGCGTCGAGA  
 CACCGCCATGGAGCTACTTGTTCGAC TAA

**TaCBFIVd-4.1 (EF028780)**

ATG GACGTGCGCCGACGCTGCCTCCAAGTCCGGCCAGCATGAGCAGGGTCACAGGACGGTGTCTGTCGGAGCCG CCGAA  
 GCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCCACCCG CTGTACCGGGCGGTGCGGGCGCGTGGCCGGGTGCG  
 GCGAGTGGGTGTGCGAGGTGCGCGTGC CGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAG  
 GTGGCGGCGCGCGCACGACGCCCGGTTGCTCGCGCTCTCGGGCGCGCCGCTGCTCAACTTCGCC GACTCCGC  
 TGGCGGATGCTGCCCCTGCTTGC CGCCGGATCGTTCCGCTTCGGCAGCGCGCGGGAGATCAAGCTTGCCGTGCGCCG  
 TCCCGTCTGTTGCGTTCAGCAGCAGCAGATTATTCTTCCAGTCGCGTGTCCAACGGTGGAGGCGCGCCGACGCCG  
 TCAACTCTCTGTTTACATGTCGTCCGTCGACTTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGC  
 TGGTCTGTAACGAGAGCTTGGCGCAGGGGATGCTCATGGCGCCCGCGGACGACAGAGCAGGCGGGAGGACGCCG  
 TCGAGACCGCGTCGAGACACCGACGCCGTTATGGAGCTATTTGTTTGAC TAA

**TaCBFIVd-B4 (EF028781)**

ATG GACGTGCGCCGACGCTGCCTCCAAGTCCGGCCACCAGGAGCAGGGTCACAGGACGGTGTCTGTCGGAGCCG CCGAA  
 TGGCGGCCCGGGGCGGACCAAGTTCCGTGAGACACGACACCCG CTGTACCGGGCGGTGCGGGCGCGTGGCCGGGTGCG  
 GCGAGTGGGTGTGCGAGGTGCGCGTGC CGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAG  
 GTGGCGGCGCGCGCACGACGCCCGGTTGCTCGCGCTCTCGGGCGCGCCGCTGCTCAACTTCGCC GACTCCGC  
 TGGCGGATGCTGCCCCTGCTTGC CGCCGGTCTGCTTTCGCTTCGGCAGCGCGCGGGAGATCAAGCTTGCCGTGCGCCG  
 TCCCGTCTGTCGCGTTCAGCTGCAGCAGAAATATTCTTCCAGCCGCGTGTCCAACGGTGGAGCCGACCGACAGCCCG  
 TCAACTCTCTGTTTACATGTCGTCCGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGC  
 TGGTCTGTAACGAGAGCTTGGCGCAGGGGATGCTCATGGCGCCCGCGGACGACAGAGCAGGCGGGAGGACGCCG  
 TCGAGACCGCGTCGAGACACCGACGCCGTTATGGAGCTATTTGTTTGAC TAA

**TaCBFIVd-26.1 (Accession)**

ATG CCGTCTGGTCAGCAGGAGCAGCGGCACAGGACGGTGAAGTTCGGAGCCG CCGAAGCGCCCCGGGGCGGACCAA  
 TCCAGGAGACGCCACCCG CTGTACCGGGCGGTGCGGGCGCGTGGCCGGGTGCG  
 GCGAGTGGGTGTGCGAGGTGCGCGTGC CGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAG  
 GTGGCGGCGCGCGCACGACGCCCGGTTGCTCGCGCTCTCGGGCGCGCCGCTGCTCAACTTCGCC GACTCCGC  
 TGGCGGATGCTGCCCCTGCTTGC CGCCGGTCTGCTTTCGCTTCGGCAGCGCGCGGGAGATCAAGCTTGCCGTGCGCCG  
 TCCCGTCTGTCGCGTTCAGCTGCAGCAGAAATATTCTTCCAGCCGCGTGTCCAACGGTGGAGCCGACCGACAGCCCG  
 TCAACTCTCTGTTTACATGTCGTCCGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGC  
 TGGTCTGTAACGAGAGCTTGGCGCAGGGGATGCTCATGGCGCCCGCGGACGACAGAGCAGGCGGGAGGACGCCG  
 TCGAGACCGCGTCGAGACACCGACGCCGTTATGGAGCTATTTGTTTGAC TAA



CTTGGCGCAGGGGATGCTGGTGGAGCCGCCGACGCCGGAGCGTGGCGAGAGCAGAGCGAGCACAGCGCGTGGC  
GGAGACGCAGACGCCGTTGTGGGC **TAA**

>TaCBFIVd-26.2a (Accession)

**ATG**CCGTCTGGTCAGGAGGAGCAGCGGCACAGGACGGTGAGGTCGGAGCCA**CCGAAGCGGCCCGCGGGGCGGACCAA**  
**GTTCCAGGAGACCGCCACCCG**CTGTACCGCGGCGTGCGCCAGCGGGGGCCCGCGGGCGGTGGGTCTGCGAGGTGC  
CGGTGCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCGTACCGCCGAGATGGCGGCGCGCGCCACGAC  
GCCGCCGTACTCGCGCTCTCCGGCCGCAAAGCCTGCCTCAACTTCGCC**GACTCCGCCTGGCGG**ATGCTGCCCCTGCT  
CGCGGCCGGCTCCTTCGGCTTCGGCAGCGCGCGGGAGATCAAGACCGCCGTCGCCGTCGCCGTCCTCGCTTCCAGC  
GGCAGCAGATCGTTCTTCCAGTCGCGTGCCCGGCGCGGAGCCGGCCGTCGCCCGAGCGGCGCGCTGTTCTCCATG  
TCGTACGGCGACTTGTGAGCTCGACGACGAGCAGTGGTTTGGCGGCATGGTCGCCGGGTCTACTACGAGAGCTT  
GGCGCAGGGGATGCTCGTGGAGCCGCCGACGCCGGAGCCTGGCGAGAGGACAGCGACCACAGCGCGTGGCGGAGA  
CGCAGACGCCGTTGTGGAGC **TAA**

>TaCBFIVd-26.2b (Accession)

**ATG**CCGTCTGGTCAGGAGGAGCAGCGGCACAGGACGGTGAGGTCGGAGCCG**CCGAAGCGGCCCGCGGGGCGGACCAA**  
**GTTCCAGGAGACCGCCACCCG**CTGTACCGCGGCGTGCGCCAGCGGGGGCCCGCGGGCGGTGGGTCTGCGAGGTGC  
CGGTGCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCGTACCGCCGAGATGGCGGCGCGCGCCACGAC  
GCCGCCGT**GCTCGCGCTCTCCGGCCGCAAAGCCTGCCTCAACTTCGCC****GACTCCGCCTGGCGG**ATGCTGCCTGTGCT  
CGCGGCCGGCTCCTTCGGCTTCGGCAGCGCGCGGGAGATCAAGACCGCCGTCGCCGTCGCCGTCCTCGCTTCCAGC  
GGCAGCAGATCGTTCTTCCAGCCGCTGCCC GGCGCGGAGCCGGCCGTCGCCCGAGCGGCGCGCTGTTCTCCATG  
TCGTACGGCGACTTGTGAGCTCGACGACGAGGAGTGGTTTGGCGGCATGGTTGCCGGGTCTACTACGAGAGCTT  
GGCGCAGGGGATGCTCGTGGAGCCGCCGACGCCGGAGCCTGGCGAGAGGACAGCGAGCACAGCGCGTGGCGGAGA  
CGCAGACGCCGTTGTGGAGC **TAA**

>TaCBFIVd-9.1a (EF028782)

**ATG**GACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGGAGCAGGGCCACCGACGGTGTCGTCGGAGCCG**CCGAA**  
**GCGGCCCGCGGGGCGGACCAAGTTCCACGAGACCGCCACCCG**CTGTACCGCGGCGTGCGGCGCGGTGGCCGCGTGC  
GGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAG  
ATGGCGGCGCGCGCGCACGACGCGCGCGTGTCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCC**GACTCCGC**  
**CTGGCGC**ATGCTGCCCCTGCTCGCGGCCGGCTCCTTCGGCTTCGGCAGCGCGAGCGAGATCAAGGCCGCCGTCGCTG  
TCGCCGTGGTTCGCTTCCAGCGGAAGCAGATTGTTCTTCCAGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGAAGCAG  
GTCCCGATCGCCGTCGCCGTGGTGGCGCTCCAGCAGAAGCAGGTTCCGGTTCGCCGTCGCCGTGGTGGCGCTCCAGCA  
GCTGCCAGTTCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATCATTCTTCCAGTCGCGTGCTGGCGC  
CCGAGTTTTACATGTCTCGCGGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGGTGCG  
TACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGGAGAGCGGCGAGCAGAG  
CGGCGTCCAGACGCCGCTATGGAGCTGCTTGTTCGAC **TAA**

>TaCBFIVd-9.1b (Accession)

**ATG**GACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGGAGCAGGGCCACCGACGGTGTCGTCGGAGCCG**CCGAA**  
**GCGGCCCGCGGGGCGGACCAAGTTCCACGAGACCGCCACCCG**CTGTACCGCGGCGTGCGGCGCGGTGGCCGCGTGC  
GGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAG  
ATGGCGGCGCGCGCGCACGACGCGCGCGTGTCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCC**GACTCCGC**  
**ATGGCGC**ATGCTGCCCCTGCTCGCGGCCGGCTCCTTCGGCTTCGGCAGCGCGAGCGAGATCAAGGCCGCCGTCGCTG  
TCGCCGTCGTCGCTTCTTCCGGAAGCAGATTGTTCTTCCAGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGAAGCAG  
GTCCCGATCGCCGTCGCCGTGGTGGCGCTCCAGCAGAAGCAGGTTCCGGTTCGCCGTCGCCGTGGTGGCGCTCCAGAA  
GCTGCCAGTTCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATCATTCTTCCAGTCGCGTGCTGGCGC  
CCGAGTTTTACATGTCTTCCGGGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGGTGCG  
TACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCCAGGCCGGAGAGCGGCGAGCAGAG  
CGGCGTCCAGACGCCGCTATGGAGCTGCTTGTTCGAC **TAA**

>TaCBFIVd-B9 (EF028783)



ATG GACGTCGCCGACATCGCCTCCCGGTCCGGCCAGCAGGAGCAGGGGCACCGAACGGTGTCTGTCGGAGCCG CCG  
 AAGCGCCCTGCGGGGGCGGACCAAGTTCCACGAGACGCGCCACCCG CTGTACCGCGCGGTGCGGCGCGGTGGCCGGGT  
 TGGGCTAGTGGGTGTGCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACTTCAACACGGCCG  
 AGATGGCGGCGCGCGCGCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCC GACTCC  
 GCCTGGCGGATGCTGCCCGTGTCTCGCGGCCGGTCTGTTCCGGTTTGACAGCGCGCGGGAGATCAAGGCCGCCGTGCG  
 CGTGGCCGTCTGTCGCTTCCAGCGGAAGCAGATCATTCCAGTCGCCGTGCGCGTCTGTCGCTCCAGCAACAGCAGG  
 TTCCGGTTCGAGTGGCCGTGGTGGCGCTCCAGCAGAAGCAGGTCCCGCTCGCCGTGCGCGTCTGTCGCTCCAGCAG  
 CTGCAGTTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCAGCAGATCATTCTCCAGTCGCGTGCCTGGC  
 GCCGAGTTTACATGTCTTCCGCGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGGT  
 CGTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGGACGAAAGAGCGGGGCCGGAGCACGGCGAGCAG  
 AGCGCGTCCAGACGCCACTATGGAGCTGCTTGTTCGAC TAA

>TaCBFIVd-D9 (EF028784)

ATG GACGTCGCCGACATCGCCTCCCGGTCTGGCCAGCAGGGGCATCGGGACACCGGACGGTGTCTGTCGGAGCCG CC  
 GAAGCGCCCCGCGGGGGCGGACCAAGTTCCACGAGACGCGCCACCCG CTGTACCGCGCGGTGCGGCGCGGTGGCCGGC  
 TCGGCGAGTGGGTGTGCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACTTCAACACGGCC  
 GAGATGGCGGCGCGCGCGCACGACGCGCGTGTGCTCGCGCTCATCGGCGCGCGCGCTGCCTCAACTTCGCC GACTC  
 CGCATGGCGCATGCTGCCCGTGTCTCGCGGCCGGTCTGTTCCGGTTTCGGCAGCGCGAGCGAGATCAAGGCCGCCGTGCG  
 CCGTGGCCGTCTGTCGCTTCCAGCGGAAGCAGATTATTCGGTTCGCCGTGCGCGTCTGTCGCTCCAGCAGCAGCAG  
 GTTCCGGTTCGCCGTGCGCGTCTGTCGCGTCCAGCAGAAGCAGGTTCGGTTCGCCGTGGCCGTCTGTCGCGTCCAGCA  
 GCTGCAGTTCCGGTTCGCCGTGGCCGTCTGTCGCGTCCAGCAGCAGCAGATTATCTTCCAGTCGCGTGCCTGGCGC  
 CGGAGTTTACATGTCTTCCGCGACCTGTTGGAGCTTGACGAGGAGCAGTGGTTTGGCGGCATGGACGCTGGGTGCG  
 TACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGGACGAAAGAGCGAGGCCGGAGCACGGCGAGCAGAG  
 CGGCGTCCAGACGCCGTATGGAGCTGCTTGTTCGAC TAA

>TaCBFIVd-A22 (EF028785)

ATG GACGTCGCCGACGCCGCTCCTCATCTGGCCAGGAGCAGCAGGGGCACAGGACGGTGTCTGTCGAGCCGCCGAA  
 GCGCCCCGCGGGGGCGGACCAAGGTGCATGAGACGCGCCACCCG CTGTACCGCGCGGTGCGGCGAGCGGGGCCGGGTGCG  
 GGCAGTGGGTGTGCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACTTTCGCCACCGCCGAG  
 ATGGCGGCGCGCGCGCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCC GACTCCGC  
 GTGGCGCATGCTGCCCGTGTCTCGCGGCCGGTCTGTTCCGGTTTCGGCAGCGCGCGGAGATCAAGGCCGCCGTGCGCG  
 TCGCCGTCTGTCGCTTTCAGAAGGAGCAGATTATCCAGTCGCCGTGCGCGTCTGTCGCGTCCAGCAAGCAGCAGATT  
 ATCCAGTCGCGTTCGCCGTGCTAGCGCTCCAGAAGCAGCAGATTCCAGTCGCCGTGCGCGTCTGTCGCGTCCAGGA  
 ACAGCAGGTCCAGTGGCTGTGCGCGTCTGTCGCGTCCATCGGCAGCAGGTTCAGTCGCGTGCCTGGCAACCTCCG  
 GCCCGGCGAGCGCTCTGTTTTACATGTCTGTCGCGACTTGTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATG  
 GAGGCCGGTCTGACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCACCGCCGGACGAAAGAGCGAGGCCGGAGGA  
 CGGCGAGCAGAGCGCGTCCAGACGCCGTTATGGAGCCAGAGCCACCTGTTCAAC TAA

>TaCBFIVd-B22a (EF028786)

ATG GACGTCGCCGACGCCGCTCCTCATCTGGCCAGGAGCAGCAGGGGCACAGGACGGTGTCTGTCGAGCCGCCGAA  
 GCGACCCGCGGGGGCGGACCAAGGTACATGAGACGCGCCACCCG CTGTACCGCGCGGTGCGGCGAGCGGGGCCGGGTGCG  
 GGCAGTGGGTGTGCGAGGTGCGCGTGGCCGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACTTCAACACCGCCGAG  
 ATGGCGGCGCGCGCGCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCGCTGCCTGCCTCAACTTCGCC GACTCCGC  
 CTGGCGCATGCTGCCCGTGTCTCGCGGCCGGTCTTTCGGTTCGGCAGCGCGCGGAGATCAAGGCCGCCGTGCGCG  
 TCGCCGTCTGTCGCTTTCAGAAGGAGCAGATTATCCAGTCGCCGTGCGCGTGGTGGCGATCCAGAAGCAGCAGATT  
 ATCCAGTCGCTGTGCGCGTCTGTCGCGATCCAGAAGCAGCAGATTATCCGTGCGCGTCTGTCGCGTCCAGAAGCAG  
 CAGATTCCAGTGGCCGTGCGCGTCTGTCGCGTCCAGGAACAGCAGGTCCAGTCGCCGTGCGCGTCTGTCGCGTTC  
 TCGGCAGCAGGTTCAGTCGCGTGCCTGGGAACCTCCGGCCCGGGCAGCGCTCTGTTTTACATGTCTGTCGCGACT  
 TGTGGAGCTGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGGTCTGACTACGCGAGCTTGGCGCAGGGGATG  
 CTCGTGGCACCGCCGGACGAAAGAGCGAGGACGGAGGACGGCGAGCAGAGCGGCGTCCAGACGCCGCTATGGAGCCA  
 CTTGTTACCT TAA

**TaCBFIVd-B22b (Accession)**

ATGACGTCGCCGACGCCGCTCCCCGTCTGGCCAGGAGCAGCAGGGGCACAGGACGGTGTCTGAGCCG**CCGAA**  
CGACCCCGCGGGGCGGACCAAGGTACATGAGACGCGCCACCCGCTGTACCGCGGCGTGCAGCAGCGGGGACGGGTCTG  
GTGGGTGTGCGAGGTGCGCGTGGCCGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAG  
ATGGCGGCGCGCGCACGACACCGCCGTGCTCGCGCTCTCCGGCCGGGCTGCCTGCCTCAACTTCGCGGACTCCGC  
**CTGGCGC**ATGCTGCCCGTGTCTCGCGCCGGGTCTTTCCGGCTTCGGCAGCGCGCGGAGATCAAGGCCGCCGTGCGCG  
TCGCCGTCTGGCTTTCCAGAAGGAGCAGATTATCCAGTCGCCGTGCGCGTGGTGGCGATCCAGAAGCAGCAGATT  
ATCCAGTCGCTGTGCGCGTCTGGCGATCCAGAAGCAGCAGATTATCCGTCGCCGTGCGCGTGGTGGCGCTCCAGAAGCAG  
CAGATTCCAGTGGCGTCCGCTCTGGCGCTCCAGGAACAGCAGGTCCAGTCGCCGTGCGCGTCTGGCGCTTCA  
TCGGCAGCAGGTTCCAGTCGCGTGGCCGGAACTCCGGCCCGGCGAGCGCTCTGTTTTACATGTCTCCAGCGACT  
TGTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATG  
CTCGTGGCACCGCCGGGCGAAAGAGCGAGGACGGAGGACGGCGAGCAGAGCGCGTCCAGACGCCGCTATGGAGCCA  
CTTGTTACCC**TAA**

**TaCBFIVd-D22 (EF028787)**

ATGACGTCGCCGACGCCGCTCCCCGTCTGGCCAGGACCAGCAGGGGCACAGGACGGTGTCTGAGCCG**CCGAA**  
CGACCCCGCGGGGCGGACCAAGGTCCATGAGACGCGCCACCCGCTGTACCGCGGCGTGCAGCAGCGGGGACGGGTCTG  
GTGGGTGTGCGAGGTGCGCGTGCCTGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAG  
ATGGCGGCGCGCGCACGACCGCCCGTGTCTCGCGCTCTCCGGCCGGGCTGCCTGCCTCAACTTCGCGGACTCCGC  
**ATGGCGC**ATGCTGCCCGTCTCTCGCGCCGGGTCTTTCCGGCTTCGGCAGCGCGCGGAGATCAAGGCCGCCGTGCGCG  
TCGCCGTCTGGCGTTCAAGAAGCAGCAGATTATCCAGTGGCCGTGCGCGTCTGGCGCTACAGAAGCAGCAGATT  
ATCCAGTGGCCGTGCGCGTCTGGCGCTCCAGAAGCAGCAGATTCCAGTCGCCGTGCGCGTCTGGCGCTCCAGGA  
AAGCAGATCCAGTCGCCGTGCGCGTCTGGCGCTCCATCGGCAGCAGGTTCCAGTCGACGACCCGGCAACCTCCG  
GTGGCGAGCGCTCTGTTTTACATGTCTCCAGCGACTTGTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATG  
ATGGCCGGGTCTGTTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCACCGCCGGACGAAAGAGCGAGGCCGGAGGA  
ATGGCGCAGAGCGCGTCCAGACGCCGCTATGGAGCCAGAGCCACTTGTTCAC**TAA**



**Figure S2 : Liste des séquences codantes de CBF chez *T. monococcum*, *E. vulgare* et *S. cereale*.**

Les régions surlignées en jaune/rouge sont les séquences signature/AP2 utilisées dans les analyses phylogénétiques. Les gènes et les numéros d'accès mis en évidence en rose ont été utilisés pour produire les figures 1 à 3. Toutes les séquences ont été traduites et analysées par HCA.

**>TmCBFII-5 (AY951947)**

ATGGACAACCTCCGGCGTGGTCTTCTATGGCGGCGCATACGCGACGGTGATGTGCGCGCCG **CCGAAGCGGCCGGCGGG**  
**GCGGACCAAGTTCCGGGAGACGCGCCACCCG**GTGTACCGCGCGTGCGCCGGCGCGGCCCGGGGGCGCTGGGTGT  
**CCGAGGTGCGCCAGCCCAACAACAAGTCCCGCATCTGGCTCGGCACCTTCGCCAGCCCGAGGCCCGCGCGCGCGCC**  
**ACGACGTGCGCGCGCTCGCGCTCCGGGGCCCGCGCGCTGCCTCAACTTCGCC**GACTCGGCGCGCTGCTCGCCGT  
 CGACCCGGCCACGCTCCGCACGCCCCAGGACATCAGAGCCGCCGAATCACGCTCGCCAGACGGCCTGCCCGCACG  
 ACGCGCCGAGGTCTCTGTGTCCGCGCGTCTGCGCCGGCGCCCGCATGGTGATCACGCAGGAGGCCGCGGCTGCG  
 CCGTACGACAGCTACGCCATGTACGGCGGCTTGGCGGACCTGGAACAGCATTCCCACTGCTACTACGACGGGATGAG  
 CGGCAGCGGCGACTGGCAGAGCATCTCACACATGAACGTCGCCGACGAAGACGGTGGCTACGGCGCAGGAGACGTGC  
 GCTCTGGAGCTACTGA

**>TmCBFIIb-18 (AY951946)**

ATGGACATGAGCCTCGAGCACTCGAGCTCTGCCTCCTCCTCGTCCACGACCGAGCGCGGGACGGCGTGGCCGTG  
 GCCC **CCGAAGCGCCCGCGGGCCCGACCAAGTTTCGGGAGACGCGGCACCC**GTGTCCCGCGCGTGCGCCCGCGCG  
 CGAACGCGCGGCGGTGGGTCTGCGAGGTGCGCGTCCCGGGGACCGCGCACGCGCTCTGGCTCGGGACGTACTTC  
 ACUGCCGAGGCGCGCGCGCGCGCACGACCGCGCATCTCATGCTGCGTGGCGGTTCGCCCGCGTGCCTCAACTT  
 CGCG **GACTCCGCGTGGCTG**CTTTCGGTGCCGCCCGCTTCTCCAACCTCTCTGATGTCCGGCGCGCGCGCTCCAGG  
 CCGTGGCGGACTTCTGCGCCGTCCGAGGCCACCGGTGCCTTCGCTGGGGCGGCTCAGGAAGTCACGTCCAGCGTG  
 ACCGTCCCGTCCGCGCGCGCGTGCAGTGTCCCTCTGTCGAGACGGCGCAGACCTCCGTGATGCCAATTTGAAGA  
 ACCGGGCGCATTAAGCATGGATATGTTGACCTCGACTGCCTGTTTGGGGAGACGACTCGGACACGTACTACTACG  
 CGAACCTTGCGCAGGCGCTGCTCATGGAGCCACCGCTAGCATGGCCACCGGGCGTACTGGGACAATGGAGACTGC  
 GCCGACGGCGGAGCCGGAGCTGATGTGCGCTCTGGAGTTACTAG

**>TmCBFIIIc-10 (AY951950)**

ATGGGCATGGGCCTTGAGATCTCGAGCTCCTCCCCCTCCTCTTCCAACGAGAACCGGTTGGTG **GCCAAAGCGCCCGG**  
**GGGGCGCACCAAGTTCCGGGAGACAAGGCACCCG**GTGTACCGCGCGTGCGCCCGCGGGCAACGCCGAACGGTGGG  
 TCTGCGAGGTGCGCGTCCCGGCAAGCGCGCGCGAGGCTCTGGCTCGGGACGTATGCCACGGCCGAGATCGCAGCG  
 UGCGCCAACGACGCGCGCATGCTCGCCCTGGGCGGCGCTCCGCCGCGCGGCTCAACTTCCCG **GACTCCGCGTGGCT**  
 GCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGCGGTCGAGGCCGTGCGGATTTGCAGC  
 GACGGGAGGCTGCCGTGGGTCCATCACCGCCACCGTCAACGAGGAGGCCTCCTGTGGCGCTCCTGCGGAGTCGTG  
 TCGGAGTCTGACGATGCCGTTCTGTCGGAGACGTGAAACCTTCTGCCCGTGGGGACTTTGCGCTGCCGGGCGGAAT  
 GGACGTGAAATGTTAGTAGGCTTGACTTGTTCCCGAAACGACTTGGGCTCGTACTACGCGAGCCTCGCGGAGG  
 CGCTGCTCATGGACCCACCGCGGTGGCGACCGGACCGCGCGTACTGGGATGACGGAGAGTTCGGCGAGGTGGCA  
 ACTGAGTTCGCGCTCTGGAGCTTAG

**>TmCBFIIIc-13 (AY951951)**

ATGGATCTTTCCAGCTCCTCCCCGTCTCGTCGGCGTCATCTCGCCCGAGCACGCGTCGGGGCGGGCGTCGCCGGC  
**CAAGCGCCCCCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCG**GTGTACCGCGCGTGCGCCCGAGGGCAACG  
 CGGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGTGGCTCTCGGCTCTGGCTCGGGACGTACCTGACGGCG  
 CAGGCGCGCGCGCGCGCAACGACGCGCGCATGCTCGCACTCGGCGGCGGCTCAGCAAGGTGCCTCAACTTCGCGGA  
**CTCGGCGTGGCTG**CTCGCCGTGCCGTCCGCGCTCTCCGACCTCGCCGACGTCCGGCGCGCGGCCCTGCAGGCCGTGCG  
 CGGATTTCCAGCGATGGGAGGCTGCCAATGGCCTGGTAACAGGACCGCGCTGAGCAGGCCCATCTAGCGCTCCT  
 GCACAATCTTCGTCTGAGTCCGACAGCGCGGATTCTGTCGGAGACGTGGAAGCTTCCGCCGATGGAGAGTTCGAGGT  
 GCTGGCTACAATGGACATCGACATGTTAGGCTTGACCTGTTCCCGAAATGGACCTGGGCTCGTACTACGTGAGCC



TCGCGGAGGCGCTGCTCATGGACCCGCCGTCGACGGCCACCATCATAGACGCGTACCGGGACAACCGTGACGGC  
GGTGCTGATGTGCCCCCTGAGGCTACTAG

>TmCBFIIId-12 (EU076381) (G3116)

ATGGACACGGGCCCCGAGCGCAACTGGAACCTCGCCGGCATCCCCGCCGTCTCGCTCGAGCAGGGGATGCCGACGTC  
GCCTGCGTGCCTGACGCGGAAGCGCCCCCGGGGGCGAACCAGTTCAAGGAGACGCGCCACCCGCTGTTCACGGCG  
TGGCGCGCGGGGCGAGCAACGGCGGGTGGGTGTGGAGGTGCGCGTGCCTGGGGAAGCGGGCGAGCGGCTCTGGCTC  
GGCAGCGACGTACCGCCGAGGGCGGCGCGCGCGCACGACGCGCCATGCTCGCGCTGTACGGCGCACCCCGCG  
TGGCGCGCTCAACTACCCCGACTCGGCGTGGCTGCTCGCCGTGCCCTCCTCCCTCTCCGACCTGGCGGACGTCCGGC  
GCGCCGCCATCGGGGCGGTGGTGGACTTCCTGCGCCGGCAGGAGCGGGCGCCAGCGCCGGCGCGCTCGCCGAGGAG  
GCCACGTGACGGGATCGCCTCCGCCGCGTGGTGCCAGACAATGCCAGCTCGTGGCGGGCGGCGTCCCACTCCCA  
CTCAGCCGCCATGCGCCAATGCGGGGTACGAGGTGCCGAGCAGATTGTGCCACGACATGTTGAGCTCCACACGT  
CGGGCGAAATGGACGCGGGCACGTACTACGCGGACCTGGCGCAGGGGCTGCTCCTGGAGCCGCCGCCGCCGCTCC  
AGCGGGGCGAGCTCGGAGCGCGGAGACGATGCGGCGCTATGGAACCACTGA

>TmCBFIIId-12 (AY951944) (DV92)

ATGGACACGGGCCCCGAGCGCAACTGGAACCTCGCCGGCATCCCCGCCGTCTCGCTCGAGCAGGGGATGCCGACGTC  
GCCTGCGTGCCTGACGCGGAAGCGCCCCCGGGGGCGAACCAGTTCAAGGAGACGCGCCACCCGCTGTTCACGGCG  
TGGCGCGCGGTGGGTGTGGAGGTGCGCGTGCCTGGGGAAGCGGGCGAGCGGCTCTGGCTCGGCACGCGACGTACCC  
GCGGAGGCGGCGCGCGCGCGCACGACGCGCCATGCTCGCGCTGTACGGCGCGACCCCGCGCGCGCTCAACTA  
CCCGACTCGGCGTGGCTGCTCGCCGTGCCCTCCTCCCTCTCCGACCTGGCGGACGTCCGGCGCGCCGCCATCGGG  
CCGTGGTGGACTTCCTGCGCCGGCAGGAGCGGGCGCCAGCGCCGGCGCCGTGCGCGAGGAGGCCACGTGACGGG  
ATCGCCTCCGCCGCGTGGCGGCCAGACAATGCCAGCTCGTGGCGGGCGGCGTCCCACTCCCACTCACAGCCGCCATG  
CGCCAATGCGGGGTACGAGGTGCCGAGCAGATTGTGCCACGACATGTTGAGCTCCACACGTGGGGCGAAATGGACG  
CGGGCACGTACTACGCGGACCTGGCGCAGGGGCTGCTCCTGGAGCCGCCGCCGCCGCTCCAGCGGGGCGAGCTCG  
GAGCACGGAGACGATGCGGCGCTATGGAACCACTGA

>TmCBFIIId-15 (EU076383) (G3116)

ATGGACATGACCGGCTCCGACAGCAGCGGAGCTCCCCGTCTCGCCGTCTCTCTCGCACCTGAAGCGTCTTG  
CGGGCGCACCAAGTTCAAGGAGACGCGGCACCCGCTGTACCGCGCGTGCCTGGCGCGCGGCGAGCGCGCGGTGG  
TGTGGAGGTGCGCGTCCCGGGCAAGCGGGCGAGCGGCTCTGGCTCGGCACGCACTCACTGCCGAGCGCGCGCG  
CGAGCGCACGACGCGCCATGCTCGGCGTATCGGCGCTCCACCCCGTGTCTCAACTTCGCCGACTCCGCTGGCT  
CTCGCCGTGCCGTCCGCGCTCTCCGACTTTCAGACGTCGCGCGCGCGGCTCTCAGCGCTGTGCGGACTTCCAAC  
GCCGGGAGGCCGCCAGCGCGCGGCCACCAGTCCCTGGCCGCCACCGTCCCCGTGATGACGGAAGCTGTAGCCAA  
TCCGCGCAGTCTCATGAAATACCGGCTCGTCTGTGGACGTCGTCGTCGTCCTACCTTCTGGCGACGGAAT  
GTTTGGGTGCCGGCCACACTGGGCTGCAACATGTTTCAGCTGGATATGTCCGGGAAATGGACCTGGACACATACT  
ATGCGTACTTTCGCGAGGGGCTTCTCTCGAGCCGCCGAACCGCGGTGCGCGCGCGTGTGGACACCGAAGGC  
GCGGAGCAGACGCGCGCTCTGGAGCTACTAA

>TmCBFIIId-16 (AY951944) (DV92)

ATGGACATGACCGGCTCGGACAGCAATGGAGCTCCTCTTCTCGCCGTCTGTCGACCTCTTCGCACCGGAAGCGCCC  
CGCCGGGCGCACCAAGTTCAAGGAGACGCGGCACCCGCTGTACCGCGCGTGCCTGGCGCGCGGGGCAACGCCGGCGCT  
GGGTGTGGAGGTGCGCGTCCCGGGCAGCGCGCGGAGCGGCTCTGGCTCGGAACGTACCTCACAGCCGACGCGGCG  
GCGCGTGCACGACGCGCGCATGCTCGGCGTGTGCGTGCCTCGCGCGGTGCCTCAATTTGCTGACTCCGCTTG  
GCTCTTGCCGTGCCGCCGCGCTCGCCGACCTCGCGGCCGTAGGCGTGCCTCGCCGCCGTGGCGGACTTCC  
AGCGGCGCATGCTCCAACAGCGCAGCCACCGTCCCCGTGATGAGGAGACCTCCGGCGCGTCCGCTCTGTCTGTCT  
GCGGACAATGCGAGCGGTTCTGTAGCGACGTCGCAACCTTGGGCGAGGGAACGTTGAGGTGCCGTCCGCGCTGGG  
CAGCGACATGTTGAGCTGGACTTGTCCGGGGAGATGGACCTGGGCACGTACTACGCGGACCTCGCGGACGGGCTGC  
TCTGGAGCCGCCCGCTGCTGGACAGCGGGGCGTGTGGGACACCGGAGACGGCGGAGCTGACTCCGGGCTCTGG  
AGCTACTGA

>TmCBFIIId-16\_ (EU076384) (G3116)



ATGGACATGACCGGCTCGGACCAGCAATGGAGCTCCTCTTCTCCTCGCCGTCGTCGACCTCTTCGCACCCGAAGCG  
 CCGCGCGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGCAGCGCGCGGGGCAACGCGCGCC  
 CTGGGTGTGCGAGGTGCGCGTCCCCGGGCAGCGCGCGAGCGGCTCTGGCTCGGAACGTACCTCACAGCCGACGCG  
 CCGCGCGTGCACGACGCGCCATGCTCGGCTGCTCGGTGCTCGCGCGGTGCCTCAATTTCTGCTGACTCCGC  
 TTGGCTCCTTGCCGTGCCGCCGCGCTCGCCGACCTCGCGCGCGTCAGGCGCGCGGCCCTCGCCGCGGTGGCGGACT  
 TCCAGCGGGCGCATGCCTCCAACAGCGCAGCCACCGTCCCCGCTGATGAGGAGACCTCCGGCGCGTCCGCTCTGTCTG  
 TCTGCGGACAATGCGAGCGGTTCTGTCAGCGATGTCGAACCTTGGGCCGAGGGAACGTTGAGGTGCCGTCCGCGCT  
 GGGCAGCGACATGTTGAGCTGGACTTGTCCGGGGAGATGGACCTGGGCACGTACTACGCGGACCTCGCGGACGGGC  
 TGCTCCTGGAGCGCCCGCTGCTGGACAGCGGGGCGTGTGGACACCGGAGACGGCGGAGCTGACTCCGGGCTC  
 TGGAGCTACTGA

>TmCBFIIId-17 (AY951945)

ATGGACATGGGCAGCGAGCAATGGAGCTCCCCGTCGACGTCGGCGTCTTCGCGGACAGCAGCGGCGCGCGCC  
 GAAGCGCCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGAGGCGTGCAGCGCGCGGGCGGAG  
 CCGGCGCGTGGGTCTGCGAGGTGCGCGTGCCTGGCAGGCGCGGGTGCAGGCTCTGGCTCGGGACCTACGTCACCGCG  
 GAGTCCGCGCGCGCGCGCACGACGCGCCATGCTCGCGCTGGGCGGCGCTCCGCGCGGTGCCTCAACTTCCCGGA  
 CTCGCGTGGCTGCTGCCGTGCCGTGCGCGCTCGCCGACCTCGCCGACGTCGGCGCGCGGGCGTCCGCCCGTGC  
 CGGGCTTCCAGCGCGGGAGGCCGCCAGCGCGCGCGACCGTTCCTGAGCAGGTTGTCGACACCTCCAGCGCG  
 GATGACGCGGGCTCCTGGTCTGTTGGCGACGCCGACGCCAGTGTGTCGGCTGCGGACGGGATGTTGAGGTGCCGGC  
 GGCAGCACTGGCCAGCGACATGTTGACTTCGAGTTGACGTGTCCTGGGTGATGGACCTGGGCTCGCCGGCGACGT  
 CGCAGCCTGGTTGTGCGGACAAAGTGTGGAGGTGCCGGCGCGCACTGGGCGCGCGGACATGTTGAGTTGAC  
 TTGGAGCTGGACATGTCGGGGAAATGAACCTGGTGGCTCTACTACGCGGATTTCGCGGAGGGGCTGCTCCTCGA  
 GCCTCCGAACCGGCCGATGCCACGGAGGCGCGTGGCGGAACGGGGACTACTGCGCGCGCGACGGCGGAGGTGACG  
 CCGCGCTCTGGAGTCAGTAG

>TmCBFIVa-2 (AY951945)

ATGGACACCGCGCGCGCTGGCCGATTTTGAAGGGCAAGAGTACAGGACGGTGTGGCCGGAGGAGGAGTACCGGAC  
 GGTGTGGTTCGGAGCCGCGAAGCGCGCGCGGGGCGGAACAAGTTGCAGGAGACGCGCCACCCGCTGTACCGCGCG  
 TCGCGCGCGTGGCGGGAAGGGCAGTGGGTGTGGGTGTGCGAGCTGCGCGTGCCTGGCTGCCGGAAGCCGGGTTTAC  
 TCCAGGATCTGGCTCGGCACCTTCGCGGACCCGAGATGGCGCGCGCGCGCACGACTCGGCGCGCTCGCGCTCTC  
 TGGCGGGGATGCGTGCCTCAACTTCGCGGACTCCGCTGGCGGATGATGCCCGTGACGCGCGCGGGTTCGTTCAAGT  
 TGGCTGCCGCTCAGGAGATCAAGGATGCCGTGGCGCTGCCCTCAAGGCATTCCAGGAGCAGCAGCCCTGCCGAC  
 CGCTCGAAGGCGCCATCCTCCACGACTCCACGTCCGAGGAGAGCGCGCGTCCATCACCTCCAACGACTTGTCTGG  
 GCTGGACGACGAGCACTGGATCGGCGCATGGACGCCGGGTCTGTAACGAACTTGGCGCAGGGCATGCTCATGG  
 AGCCGCCGGCGCGCGAGCGTGGAGGGAGGACCGCAACAGGACGACGGCGTCGACACGTGCTGTGGAGCTACTGG  
 TTGGACGGATTGATGTGTAAAGTTGTAG

>TmCBFIVc-14 (EU076382) (G3116)

ATGGACGCGCGCGACGCCCTCCCCGTGTGATGGGCACAGGACGGTGTGGTTCGGAGCCGCGAAGCGGCGGCGCG  
 CCGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGCAGCGACGGGCGCGCGCGCGGTGGGTGT  
 TCGAGGTGCGGTGCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCACCAACCGCGAGATGGCAGCGCG  
 GCGCAGCAGCGCGCGGTTCTCGCGCTCTCCGGCGCGCGGCTTGTCTCAACTTCGCGGACTCCGCTGGCGGATGCT  
 CCCCCTCCTCGCGGCGCGTTTACGACCCGCCAAGGAGATCAAGGATGCCGTGCGCGTCCGCGTCTGGCGTTCCAAA  
 GGCAGCACCCGCTCGGTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCTCG  
 CCGGCGCGGAGCGCCCTGTCCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCGTG  
 CTACTCGGAGTTTATGGAGTCGCCGACACAGACCGTGGCGGGAAGACTTTGAGCTCGGTGGCGTCGAGACACCGC  
 CATGGAGCTACTTGTTCGACTTAA

>TmCBFIVd-4 (AY951945)

ATGCCGTCTGGTCAGGAGGAGCAGCGGCACAGGACGGTGAAGTTCGGAGCCGCGAAGCGGCGGCGGGGCGGACCAA  
 GTTCCAGGAGACGCGCCACCCGCTGTACCGCGCGTGCAGCGCGGGGCGCGCGGGCGGTGGGTCTGCGAGGTGC  
 CCGTGTCTGGGATGAGGGGATCCAGGCTCTGGCTCGGCACCTTCGTACCGCGGAGATGGCGGCGCGCGCCACGAC



CCGCGCGTGGCTCGCGCTCTCCGGCCGCAAGGCTGCTCAACTTCGCGGACTCCGCTGGCGGATGCTGCCCGT  
GCTCGCGGCCGGCTCCTTCGGCTTCGGCAGCGCGCGGGAGATCAAGACCGCCGTCGCCGTCGCCGTCCTCGCGTTCC  
AGCGGCAGCAGATCGTTCTTCCAGTCGCGTGCCCGGCAGCGGAGCCGGCCGTCGCCCCGAGCGGCGCGCTGTTCTCC  
ATGTCGTAGGCGACTTGTCTGGAGCTCGACGACGAGCAGTGGTTTGGCGGCATGGTTGCCGGGTCTGTAACGAGAG  
CTTGGCGCAGGGGATGCTCGTGGAGCCGCCGAGCGCGGAGCCTGGCGAGAGGACAGCGAGCACAGCGGCGTGGCGG  
AGACGCAGACGCCGTTGTGGAGCTAA

>HvCBFI-1 (AY785837) (Dicktoo)

ATGGACGTTGGTGCCCTCAGCAGCGACTACTCGTCGGGGACGCCGTCCCCGGTGGGCGCGGACGGCGGCAACAGTGA  
GGGCTTCTCGACGTACATGACGGTGTCTTCGGCTCCGCGAAGCGCGCGCGGGGCGGACCAAGTTCAAGGAGACGC  
GGCACCCGCTCTACAAGGGCGTGGCGCGGAGGAACCCCGGGAGGTGGGTCTGCGAGGTGCGGGAGCCGCACAGCAAG  
TAGAGGATATGGCTCGGCACGTTTCGAGACCGCAGAGATGGCGGCGCGCGCGCACGAGTGGCCGCGCTGGCGCTGCG  
GGCGCGCGCGCTGCTCAACTTCGCGGACTCTCTCTGTCGGCTCCGGGTCCCGGCCGTGGGTGCCAGCCCTGATG  
AGATACGGCGGGCGCGGTAGAGGCGGCTGAGGCATTCTGCCGCAACCGACAGAGCAATGCGCCCGCGGAGGAG  
GTAGCCGCTGCACCAACGATGCAGTTCGCCGGTGTATCCGTACTACGGGATGGACGATGGTATGGACTTCGGGATGCA  
GGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCTCCGTGGTAGGTCCGTCAGCGACTGCCGAGACG  
CGACGATGACGGTGAGGTACGTTGTGGAGCTACTGA

>HVCBFI-1\_ (AY785838) (Strider)

ATGGACGTTGGTGCCCTCAGCAGCGACTACTCGTCGGGGACGCCGTCCCCGGTGGGCGCGGACGGCGGCAACAGTGA  
GGGCTTCTCGACGTACATGACGGTGTCTTCGGCTCCGCGAAGCGCGCGCGGGGCGGACCAAGTTCAAGGAGACGC  
GGCACCCGCTCTACAAGGGCGTGGCGCGGAGGAACCCCGGGAGGTGGGTCTGCGAGGTGCGGGAGCCGCACAGCAAA  
TAGAGGATATGGCTCGGCACGTTTCGAGACCGCAGAGATGGCGGCGCGCGCGCACGAGTGGCCGCGCTGGCGCTGCG  
GGCGCGCGCGCTGCTCAACTTCGCGGACTCTCTCTGTCGGCTCCGGGTCCCGGCCGTGGGTGCCAGCCCTGATG  
AGATACGGCGGGCGCGGTAGAGGCGGCTGAGGCATTCTGCCGCAACCGACAGAGCAATGCGCCCGCGGAGGAG  
GTAGCCGCTGCACCAACGATGCAGTTCGCCGGTGTATCCGTACTACGGGATGGACGATGGTATGGACTTCGGGATGCA  
GGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCTCCGTGGTAGGTCCGTCAGCGACTGCCGAGACG  
CGACGATGACGGTGAGGTACGTTGTGGAGCTACTGA

>HvCBFI-11 (AY785890) (Dicktoo)

ATGGAGTGGGCGTGCTGCGGACGGCTACTCGTCGTGCGGAACGAGTCCCCGGCGGCGCGGACGGGGAGGAGGG  
TTCGTACATGACGGTGTCTTCGGCTCCGCGAAGCGCGCGGGCGGGAGGACCAAGTTCAAGGAGACGAGGCACCCG  
CTAAAGGGGTGCGCAGCAGGAACCCCGGGCGGTGGGTCTGCGAGGTGCGCGAGCCGCAAGGGAAGCAGAGGCTA  
TTCCTCGSCACCTTCGACACCGCCGAGATGGCGGCGCGCGCGCACGAGTCCCGCCATGGCGCTCCGCGGCGCGCG  
TGGTGCCTCAACTTCGCGGACTCGCGCGGAGGCTCCCCGTGCCACCGCAGGGGGCGGGCCACGACGAGATACGAC  
GAGCCGCTGTGCAAGCGGCCGAAGTGTTCGCCCGGCGCCTGGGCAGCGCAATGCAGCTACCGTGGCGGCAGCGACT  
GCTCCGCCGGTGGCCTTGGGAAACGCCGAGCTCGTCGAGACTCTCCTTACTACCCCATGGACGGGTTAGAATCCGA  
AATGCAGGGCTATCTTGACATGGCGCACGGCATGCTCATCGAGCCACCACCAATGGCGTGGCCGTGACGTGGATCG  
AGGAGGACTACGACTGCGAGATCAGCCTGTGGAAGTACTGA

>HvCBFII-5 (AY785855) (Dicktoo)

ATGGACCACTGCGCGTGGGCCTCTATGGCGAGTACGCGACGGTGACGTCGGCGCCGCCAAGCGGCCGCGGGGCG  
GACCAAGTTCAAGGAGACCGCGCACCCGCTGTACCGGGGCGTGGCGGCGCGCGCGCCGAGGGCGGTGGGTCTGCG  
AGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTTGCCACGCCGAGGCGCGCGCGCGCGCCAC  
GAGCTCGCGCGCTCGCGCTCGGGGGCGCGCGCGCTGCCTCAACTTCGCGGACTCCGCGGCGCTGCTCCGCGTCTGA  
CCCGGCCACGCTCCGCACGCCCGAGGACATCCGAGCAGCGCCATGGCGCTCGCCAGGCGCGCTGCCCCACGACG  
CCGCGTCTCTCTGCCCCGCGCTGAAGCGGCGCTCCGCGCGCGCGCCAGCGATGGTGATGATGCAGGAGCCCGCG  
CCGGTGCCGTACGACAGCTATGCCACGGCCCTGTACGGCGACTTGACGGACCTGGACATGCATTCTACTACTGCTA  
GACGGGATGAGCGGCGGCGGCGACTGGCAGAGCATCTCGCGTATGGACGGAGCCGACGAAGACGGCATCTACGGCG  
AGGAGACGTCGCGCTCTGGAGCTACTGA

>HvCBFII-5 (AY785858) (Optic)



ATGGACCACTGCGGCGTGGGCTCTATGGCGAGTACGCGACGGTGACGTCGGCGCCG**CCCAAGCGGCCGGCGGG**  
**GCGGACCAAGTTCAGGGAGACGCGGCACCCG**GTGTACCGGGGCGTGCGGCGGCGCGCGCGCGCGGGCGGCTGGGTCT  
 GCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTTGCCACGCGCGAGGCCGCGCGCGCGCG  
 CACGACGTCGCGCGCTCGCGCTCCGGGGCGCGCGCGCTGCCTCAACTTCGCG**GACTCCGCGGCCCTG**CTCCGCGT  
 CGACCCGGCCACGCTCCGCACTCCGAGGACATCCGAGCAGCCGCCATGGCGCTCGCCAGGCCGCTGCCCCACG  
 ACGCCGCTCTCTCTGCCCCGCCCTGAAGGCGGCTCTGCGCCGGCGCCAGCGATGGTGATGGTGACGAGGCC  
 GCGGCGGCGCGTACGACAGCTATGCCACGGCCCTGTACGGCAGCTTGACGGACCTGGACATGCATTCTACTACTG  
 CTACGACGGGATGAGCGCGGCGGCGACTGGCAGAGCATCTCGCGTATGGACGGAGCCGACGAAGACGGCATCTACG  
 GCGCAGGAGACGTCGCGCTCTGGAGCTACT**TGA**

>HvCBFII-5 (AY785857) (88Ab536)

ATGGACCACTGCGGCGTGGGCTCTATGGCGAGTACGCGACGGTGACGTCGGCGCCG**CCCAAGCGGCCGGCGGGGCG**  
**GACCAAGTTCAGGGAGACGCGGCACCCG**GTGTACCGGGGCGTGCGGCGGCGCGCGCGCGCGGGCGGCTGGGTCTGCG  
 AGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTTGCCACGCGCGAGGCCGCGCGCGCGCGCGCG  
 GACGTCGCGCGCTCGCGCTCCGGGGCGCGCGCGCTGCCTCAACTTCGCG**GACTCCGCGGCCCTG**CTCCGCGTCTGA  
 CCCGGCCACGCTCCGACGCGCCGAGGACATCCGAGCAGCCGCCATGGCGCTCGCCAGGCCGCTGCCCCACGACG  
 CCGCGTCTCTCTGCCCCGCCCTGAAGGCGGCTCCGCTCCGGCGCCAGCGATGGTGATGGTGACGAGGCCGCGC  
 GCGGCGCGTACGACAGCTATGCCACGGCCCTGTACGGCAGCTTGACGGACCTGGACATGCATTCTACTACTGCTA  
 CGACGGGATGAGCGCGGCGGCGACTGGCAGAGCATCCCGGTATGGACGGAGCCGACGAAGACGGCATCTACGGCG  
 CAGGAGACGTCGCGCTCTGGAGCTACT**TGA**

>HvCBFIIIa-6 (AY785860) (Dicktoo)

ATGTGTACAGATCAAGAAGGAGATGAGCGGCGAGTCGGGCTCTCCCTGCAGCGGGGAGAACTACTACTACTCGCCCTC  
 GACGTCGCGGAGCATCAGCAGGCGAAGCAGCAGGCGGCTGGACGTCGGCGCCG**GCGAAGCGGCCGGCGGGGCGGA**  
**CCAAGTTCAGGGAGACGCGCCACCCG**GTGTACCGGGGCGTGCGGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
 GTGCGCGTGCCCGGAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGACACCGCGAGGCCGCGCGCGCGCGGAA  
 CGACGCGGCCATGCTCGCGCTCGCCGCGGGGGCGCGGGCTGCCTCAACTTCGCG**GACTCCGCGGAGCTA**CTCGCCG  
 TGCCGGCTGCCTCTCGTACCGCAGCTCGACGAGGTCCGCCACGCCGTCGTGGAGGCCGTCGAGGACTTGCTGCGG  
 CGCGAGGCGCACGAGGAGCAGCGCTCTCGGTCTCGGCACCTCTCGTCCGCGCCCTCTCCATCACCGACGA  
 CGACTCGTCTCTTCCCCGGCGGACGAGGGCTCGCCGTTTCGAGCTGGACGTCCTGAGCGACATGGGCTGGGACCTTT  
 ACTACGCGAGCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGCCGCGGCGCTCGGCGATTACGGCGAG  
 GCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCT**TAG**

>HvCBFIIIa-6 (EU332011) (OWB-D)

ATGTGTACAGATCAAGAAGGAGATGAGCGGCGAGTCGGGCTCTCCCTGCAGCGGGGAGAACTACTACTACTCGCCCTC  
 GACGTCGCGGAGCATCAGCAGGCGAAGCAGCAGGCGGCTGGACGTCGGCGCCG**GCGAAGCGGCCGGCGGGGCGGA**  
**CCAAGTTCAGGGAGACGCGCCACCCG**GTGTACCGGGGCGTGCGGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
 GTGCGCGTGCCCGGAGGCGCGGGAGCAGGCTCTGGCTCGGTACCTTCGACACCGCGAGGCCGCGCGCGCGCGGAA  
 CGACGCGGCCATGCTCGCGCTCGCCGCGGGGGCGCGGGCTGCCTCAACTTCGCG**GACTCCGCGGAGCTA**CTCGCCG  
 TGCCGGCTGCCTCTCGTACCGCAGCTCGACGAGGTCCGCCACGCCGTCGTGGAGGCCGTCGAGGACTTGCTGCGG  
 CGCGAGGCGCACGAGGAGCAGCGCTCTCGGTCTCGGCACCTCTCGTCCGCGCCCTCTCCATCACAGACGA  
 CGACTCGTCTCTTCCCCGGCGGACGAGGGCTCGCCGTTTCGAGCTGGACGTCCTGAGCGACATGGGCTGGGACCTTT  
 ACTACGCGAGCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGCCGCGGCGCTCGGCGATTACGGCGAG  
 GCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCT**TAG**

>HvCBFIIIa-6 (EU332012) (OWB-R)

ATGTGTACAGATCAAGAAGGAGATGAGCGGCGAGTCGGGCTCTCCCTGCAGCGGGGAGAACTACTACTACTCGCCCTC  
 GACGTCGCGGAGCATCAGCAGGCGAAGCAGCAGGCGGCTGGACGTCGGCGCCG**GCGAAGCGGCCGGCGGGGCGGA**  
**CCAAGTTCAGGGAGACGCGCCACCCG**GTGTACCGGGGCGTGCGGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
 GTGCGCGTGCCCGGAGGCGCGGGAGCAGGCTCTGGCTCGGTACCTTCGACACCGCGAGGCCGCGCGCGCGCGGAA  
 CGACGCGGCCATGCTCGCGCTCGCCGCGGGGGCGCGGGCTGCCTCAACTTCGCG**GACTCCGCGGAGCTA**CTCGCCG  
 TGCCGGCTGCCTCTCGTACCGCAGCTCGACGAGGTCCGCCACGCCGTCGTGGAGGCCGTCGAGGACTTGCTGCGG  
 CGCGAGGCGCACGAGGAGCAGCGCTCTCGGTCTCGGCACCTCTCGTCCGCGCCCTCTCCATCACAGACGA  
 CGACTCGTCTCTTCCCCGGCGGACGAGGGCTCGCCGTTTCGAGCTGGACGTCCTGAGCGACATGGGCTGGGACCTTT  
 ACTACGCGAGCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGCCGCGGCGCTCGGCGATTACGGCGAG  
 GCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCT**TAG**



CGCGAGGCGCACGCGGAGGACGACGCGCTCTCGGTCTCTTGACCTCCTCGTCCGCGCCCTCCTCCCTCACC GA  
 CGACGAGTCGTCTCTTCCCCGGCGGCCGAGGGCTCGCCGTTTCAGCTGGACGTCTGAGCGACATGGGATGGGACC  
 TTTACTACGCGAGCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGGCCGCGGCGCTCGCCGATTACGGC  
 GAGGCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCTAG

>HvCBFIIIa-6 (AY785862) (88Ab536)

ATGTGTCAGATCAAGAAGGAGATGAGCGGCGAGTCGGGCTCTCCCTGCAGCGGGGAGAACTACTACTACTCGCCCTC  
 GACGTGCGCCGAGCATCAGCAGGCGAAGCAGCAGGCGGCGTGGACGTGCGCGCCGCGGAAGCGGCCGGCGGGCGGA  
 CCAAGTTCAGGGAGACGCGCCATCCG GTGTACCGCGGCGTGC GGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
 GTGCGCGTGC CGGCGAGGCGCGGGAGCAGGCTCTGGCTCGGTACCTTCGACACCGCGGAGGCCCGCGCGCGCGGAA  
 CGACGCGCCCATGCTCGCGCTCGCCGCGGGGGCGCGGGCTGCCTCAACTTCGCC GACTCCGCGGAGCTACTCGCCG  
 TGCCGGCTGCCTCCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTGCTGGAGGCCGTGAGGACTTGTGCGG  
 CGCGAGGCGCACGCAGAGGACGACGCGCTCTCGGTCTCTGGGCACCTCCTCGTCCGCGCCCTCCTCCATCACCAGCA  
 CGACTCGTCTCTTCCCCGGCGGACGAGGGCTCGCCGTTTCAGCTGGACGTCTGAGCGACATGGGCTGGGACCTTT  
 ACTACGCGAGCCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGGCCGCGGCGCTCGGCGATTACGGCGAG  
 GCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCTAG

>HvCBFIIIa-6 (EU332009) (Scarlett)

ATGTGTCAGATCAAGAAGGAGATGAGCGGCGAGTCGGGCTCTCCCTGCAGCGGGGAGAACTACTACTACTCGCCCTC  
 GACGTGCGCCGAGCATCAGCAGGCGAAGCAGCAGGCGGCGTGGACGTGCGCGCCGCGGAAGCGGCCGGCGGGCGGA  
 CCAAGTTCAGGGAGACGCGCCACCCG GTGTACCGCGGCGTGC GGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
 GTGCGCGTGC CGGCGAGGCGCGGGAGCAGGCTCTGGCTCGGTACCTTCGACACCGCGGAGGCCCGCGCGCGCGGAA  
 CGACGCGCCCATGCTCGCGCTCGCCGCGGGGGCGCGGGCTGCCTCAACTTCGCC GACTCCGCGGAGCTACTCGCCG  
 TGCCGGCTGCCTCCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTGCTGGAGGCCGTGAGGACTTGTGCGG  
 CGCGAGGCGCACGCAGAGGACGACGCGCTCTCGGTCTCTTGACCTCCTCGTCCGCGCCCTCCTCCCTCACCAGCA  
 CGAGTCGTCTCTTCCCCGGCGGCGGAGGGCTCGCCGTTTCAGCTGGACGTCTGAGCGACATGGGATGGGACCTTT  
 ACTACGCGAGCCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGGCCGCGGCGCTCGGCGATTACGGCGAG  
 GCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCTAG

>HvCBFIIIa-6 (EU332008) (Collins)

ATGTGTCAGATCAAGAAGGAGATGAGCGGCGAGTCGGGCTCTCCCTGCAGCGGGGCGAACTACTACTACTCGCCCTC  
 GACGTGCGCCGAGCATCAGCAGGCGAAGCAGCAGGCGGCGTGGACGTGCGCGCCGCGGAAGCGGCCGGCGGGCGGA  
 CCAAGTTCAGGGAGACGCGCCACCCG GTGTACCGCGGCGTGC GGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAG  
 GTGCGCGTGC CGGCGAGGCGCGGGAGCAGGCTCTGGCTCGGTACCTTCGACACCGCGGAGGCCCGCGCGCGCGGAA  
 CGACGCGCCCATGCTCGCGCTCGCCGCGGGGGCGCGGGCTGCCTCAACTTCGCC GACTCCGCGGAGCTACTCGCCG  
 TGCCGGCTGCCTCCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTGCTGGAGGCCGTGAGGACTTGTGCGG  
 CGCGAGGCGCACGCAGAGGACGACGCGCTCTCGGTCTCTGGGCACCTCCTCGTCCGCGCCCTCCTCCATCACCAGCA  
 CGACTCGTCTCTTCCCCGGCGGACGAGGGCTCGCCGTTTCAGCTGGACGTCTGAGCGACATGGGCTGGGACCTTT  
 ACTACGCGAGCCTGGCGCAGGGGATGCTCATGGCGCCGCTGCTTCTTGGCCGCGGCGCTCGGCGATTACGGCGAG  
 GCTCACCTCGCCGACGTGCCACTCTGGAGCTACCAGAGCTAG

>HvCBFIIIc-3 (AY785845) (Dicktoo)

ATGGACATGGGCTTGAGGTCTCGAGCTCTCCCCCTCCTCGCCGGTGTCTCTCGCCCGAGCACGCGGCGAG  
 GCGGGCGTGC CGGCGCAAGCGCCCCGCTGGGCGCACCAAGTTCGGGAGACGCGGCCACCCG GTGTACCGCGGCGTGC  
 GCGCGGGGCAACACCGAAGCGGTGGGTCTGCGAGGTGCGCGTCCCGGCAAGCGCGGTGCTCGGCTCTGGCTCGGG  
 GGTACCGCACGGCTGAGGTGCTGCGCGCGCAACGACGCTGCCATGCTCGCCCTGGGCGGCGCTCGCCCGCGTG  
 CCAACTTCGCC GACTCCGCGTGGCTGCTCGCCGTGCCGTCCGCCCTGTCCGATCTCGCAGACGTCCGCGCGCGG  
 CTGTGAGGCCGTGCGGATTTCCAGCGACGGGAGGCTGCCGATGGCTCCCTCGCCATCGCTGTCCCTAAGGAGGCC  
 TCCTCTGGCGCTCCTTCACTATCTCGTCTGCTGGGTCCGACAGTGCCGTTTCGACGGGGACGTGGAACCTTCGG  
 CAATGGAGAGTTCGAGGGACCGGTTGTAATGGACAGTGAAATGTTACGGCTTGACTGTTCGGGAAATGGACCTGG  
 GCTCGTACTACATGAGCCTCGCGGAGGCGCTGCTCATGGACCGCGCCTACAGCGACCATCATCCACGCGTACGAA  
 GACAACGGCGACGGGGAGCTGATGTCGGCTCTGGAGCTATAGTGTGATATGTGA



## &gt;HvCBFIIIc-3 (EU331983) (Kompoltikorai)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCCTCCTCCTCGTCGGTGTCTGCTCCTCGCCGAGCACGCGGCGAG  
 GCGGGCGTCGCCG**GCCAAGCGCCCCGCTGGGCGCACCAAGTTCCGGGAGACGCGGCACCCG**GTGTACCGCGCGTGC  
 GCGCGCGGGGCAACACCGAACGGTGGGAGCGCGGTGCTCGGCTCTGGCTCGGG  
 ACGTACGCCACGGCTGAGGTGCG**CGCGCGCGCAACGACGCTGCCATGCTCGCCCTGGGCGGCGGCTCCGCCCGCTG**  
 CCTCAACTTCGCG**GACTCCGCGTGGCTGCTCGCGGTGCCCTGTCCGATCTCGCAGACGTCCGGCGCGCGG**  
 CTGTCGAGGCCGTGCGGATTTCCAGCGAGGGAGGCTGCCATGGCTCCCTCGCCATCGCTGTCTGCTAAGGAGGCC  
 TCCTCTGGCGCTCCTTCACTATCTCCGTGCTCTGGGTCCGACAGTGCCGGTTCGACGGGGACGTGCGAACCTTCCGC  
 CAATGGAGTGTTGAGGGGCCCCTTGTAAATGGACAGTGAAATGTTTACGGCTTGACTTGTTCGCGAAATGGACCTGG  
 GCTCGTACTACATGAGCCTCGCGGAGGCGCTGCTCATGGACCCGCCCTACAGCGACCATCATCCACGCGTACGAA  
 GACAACGGCGACGGGGGAGCTGATGTCCGGCTCTGGAGCTATAGTGTGATAT**GTGA**

## &gt;HvCBFIIIc-3 (EU331985) (OWB-R)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCCTCCTCCTCGTCGGTGTCTGCTCCTCGCCGAGCACGCGGCGAG  
 GCGGGCGTCGCCG**GCGAAGCGCCCCGCTGGGCGCACCAAGTTCCGGGAGACGCGGCACCCG**GTGTACCGCGCGTGC  
 GCGCGCGGGGCAACACCGAACGGTGGGTCTGCGAGGTGCGCGTCCCGCGCAAGCGCGGTGCTCGGCTCTGGCTCGGG  
 ACGTACGCCACGGCTGAGGTGCG**CGCGCGCGCAACGACGCTGCCATGCTCGCCCTGGGCGGCGGCTCCGCCCGCTG**  
 CCTCAACTTCGCG**GATTCGCGTGGCTGCTCGCGGTGCCGTCCGCCCTGTCCGATCTCGCAGACGTCCGGCGCGCGG**  
 CTGTCGAGGCCGTGCGGATTTCCAGCGACGGGAGGCTGCCGATGGCTCCCTCGCCATCGCTGTCCCTAAGGAGGCC  
 TCCTCTGGCGCTCCTTCACTATCTCCGTGCTCTGGGTCCGACAGTGCCGGTTCGACGGGGACGTGCGAACCTTCCGC  
 CAATGGAGTGTTGAGGGGCCCCTTGTAAATGGACAGTGAAATGTTTACGGCTTGACTTGTTCGCGAAATGGACCTGG  
 GCTCGTACTACATGAGCCTCGCGGAGGCGCTGCTCATGGACCCGCCCTACAGCGACCATCATCCACGCGTACGAA  
 GACAACGGCGACGGGGGAGCTGATGTCCGGCTCTGGAGCTATAGTGTGATAT**GTGA**

## &gt;HvCBFIIIc-3 (EU331982) (Scarlett)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCCTCCTCCTCGTCGGTATCTGCTCCTCGCCGAGCACGCGGCGAG  
 GCGGGCGTCGCCG**GCGAAGCGCCCCGCTGGGCGCACCAAGTTCTCGAGACAAGGCACCCG**GTGTACCGCGCGTGC  
 GCGCGCGGGGCAACACCGAACGGTGGGTCTGCGAGGTGCGCGTCCCGCGCAAGCGCGGTGCTCGGCTCTGGCTCGGG  
 ACGTACGCCACGGCTGAGGTGCG**CGCGCGCGCAACGACGCTGCCATGCTCGCCCTGGGCGGCGGCTCCGCCCGCTG**  
 CCTCAACTTCGCG**GACTCCGCGTGGCTGCTCGCGGTGCCGTCCGCCCTGTCCGATCTCGCAGACGTCCGGCGCGCGG**  
 CTGTCGAGGCCGTGCGGATTTCCAGCGACGGGAGGCTGCCGATGGCTCCCTCGCCATCGCTGTCTGCTAAGGAGGCC  
 TCCTCTGGCGCTCCTTCACTATCTCCGTGCTCTGGGTCCGACAGTGCCGGTTCGACGGGGACGTGCGAACCTTCCGC  
 CAATGGAGTGTTGAGGGGCCCCTTGTAAATGGACAGTGAAATGTTTACGGCTTGACTTGTTCGCGAAATGGACCTGG  
 GCTCGTACTACATGAGCCTCGCGGAGGCGCTGCTCATGGACCCGCCCTACAGCGACAATCATCCACGCGTACGAA  
 GACAACGGCGACGGGGGAGCTGATGTCCGGCTCTGGAGCTATAGTGTGATAT**GTGA**

## &gt;HvCBFIIIc-3 (EU331981) (Collins)

ATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCCTCCTCCTCGTCGGTGTCTGCTCCTCGCCGAGCACGCGGCGAG  
 GCGGGCGTCGCCG**GCCAAGCGCCCCGCTGGGCGCACCAAGTTCCGGGAGACGCGGCACCCG**GTGTACCGCGCGTGC  
 GCGCGCGGGGCAACACCGAACGGTGGGTCTGCGAGGTGCGCGTCCCGCGCAAGCGCGGTGCTCGGCTCTGGCTCGGG  
 ACGTACGCCACGGCTGAGGTGCG**CGCGCGCGCAACGACGCTGCCATGCTCGCCCTGGGCGGCGGCTCCGCCCGCTG**  
 CCTCAACTTCGCG**GACTCCGCGTGGCTGCTCGCGGTGCCGTCCGCCCTGTCCGATCTCGCAGACGTCCGGCGCGCGG**  
 CTGTCGAGGCCGTGCGGATTTCCAGCGACGGGAGGCTGCCGATGGCTCCCTCTCCATCGCTGTCTGCTAAGGAGGCC  
 TCCTCTGGCGCTCCTTCACTATCTCCGTGCTCTGGGTCCGACAGTGCCGGTTCGACGGGGACGTGCGAACCTTCCGC  
 CAATGGAGTGTTGAGGGGCCCCTTGTAAATGGACAGTGAAATGTTTACGGCTTGACTTGTTCGCGAAATGGACCTGG  
 GCTCGTACTACATGAGCCTCGCGGAGGCGCTGCTCATGGACCCGCCCTACAGCGACAATCATCCACGCGTACGAA  
 GACAACGGCGACGGGGGAGCTGATGTCCGGCTCTGGAGCTATAGTGTGATAT**GTGA**

## &gt;HvCBFIIIc-8A (AY785868) (Dicktoo)

CCGAAGCGCCCGCGGGGCGCACTAAGCTCCGGGAGACCCGGCACCTGAGTACCGTAGCGTGGGTGCGGGGGCA  
 GCGGACGGTGGGTCTGCGAGGTGCGGTCCGAGGCAAGCGCGCTCGGCTCTCGCTCGGGCCGTACCTCACCGG



CGAGGCGCGCGCCACGCCAACGAGCGCGCCATGCTCGTGTCTCGGCGCGCGCTCCGCCAGGTGCCTCAACTTC  
 CGGACTCCGCGTGGTGCTCGTTGTGTCTCCGTGCTCTCCGATCGCGCCGACGTCCGACGCGCGGCCCTCCAGGCC  
 GTAGCGAATTTCCAGCGATGGGAGGCTGCCGATGGCCTGATAACCAGGACCGTCGCCGAGAAGGCCCATCTAGCGCT  
 CCTACACAATTGTCTCGTCCGAGTCCGACAATGCTGATTATCGGAGACGTTCGGAACCTTCCGCCGATGGAGAGTTCAA  
 GTTGGCGGTTCGCGACGGACAGCGATATGTTGAGGCTTGACATATTTCCCGGAAATGGACCTAGGCTCATACTACGTGA  
 GCCTCGCGGAGGCGTGTCTATGGACACACCGTCGACGGCCACCATCATCGACACGTACCGGGACAACGGCGATGGA  
 GGAGCTGATGTCCCGCTCTGGAGCTACTAG

>HvCBFIIIc-8B (AY785873) (88Ab536)

TCGAAGCGCCCCGTGGGGCGCACTAAGCTCTGGGAGACCCGGCACCCGAGTACCGTAGCATGCGGCGCGGGGCAAA  
 CGCGGACGGTGGGTCTGCGAGGTGCGCGTCCCGAGGCAAGCGCGGCGCTCGGCTCTCGCTCGGGCCGTACCTCACC  
 CGGAGGCGCGCGCCACGCCAACGAGCGCGCCATGCTCGCGCTCGGCGCGCGCTCCGCCAGGTGCCTCAACTTCGCGG  
 ACTCCGCTGTGGTGCTCGTCTGTGTCTCCGTGCTCTCCGATCGCGCCGACGTCCGACGCGCGGCCCTCCAGGCCGTA  
 GCGAATTTCCAGCGATGGGGGGTTGCCGATGGCCTGATAACCAGGACGGTCGCCGAGAAGGCCCATCTAGCGCTCCT  
 GCACAATTGTCTCGTCCGAGTCCGACAATGCTGATTATCAGAGACGTTCGGAACCTTCCGCCGATGGAGAGTTCAAGTT  
 GCCGGTTCGCGACGGACAGCGATATGTTGAGGCTTGACATATTTCCCGGAAATGGACCTAGGCTCGTACTACCTAGACC  
 TCGCGGAGGCGTGTCTGTGGACACGCCGTGACGGCCACCATCATCGACACGTACCGGGACAACGGCGATGGTGA  
 GCTGATGTCCCGCTCTGGAGCTACTAG

>HvCBFIIIc-8C (AY785874) (Morex)

CCGAAGCGCCACGCGGGGCGCACTAAGCTTCGGGAGACCCGACACCTGTGTACCGTAGCTTGGCGCGCAGGGTCAA  
 CGCGGACGGTGGGTCTGCGAGGTGTGCGTCCCGGCAAGCGCGGCGCTCGGCTCTGGCTCGGGTCTGACCTACCGT  
 CGAGGCGCGCTCCTGTGCCAACGAGGCGCGCATGCTCGGGCTCAGTGGCGCTCCGCCAGGTGCCACAACCTTCGTGG  
 ACTCCGCGTGGTGCTCGTCTGTGTCTCCGCGCTCTCCGACCACGCCGACGTCCGGCGCGCGGCCCTCCAGGCCGTA  
 GCGAGTTTCCAGCGACGGGAGGCTGCCGATGGCCTGATAACCAGGACCGTCGCCGAGGAGGCCCATCTATTGCTCC  
 TGCACAATTGTCTCGTCCGAGTTCGAACAATGCTGATTATCAGACAGACGTGGGAAGCTTCCGCCGATGGTGAAGTTCAAGT  
 TGCCGGCCGCGACGGACAACGATATGTTGAGGCTTGACATATTTCTGAAATGGACCTAGGCTCGTACTATGTGAGC  
 CTCGCGGAGGCGTGTCTATGGACCCGCGTTGACGGCCACCATCATAGACACGTACCGGGACAACAGCGATGGCGG  
 AGCTGAGGTCCCGCTCTGGAGCTACTAG

>HvCBFIIIc-10A (AY785882) (Dicktoo)

ATGACATGGGCCTTGAGGTCTCGAGCTCTCCCCCTCTCTTCCAACGAGAACGCGTCGGGACGGTCTGTCGACGGC  
 CAAGCGCCCCGGCGGGGCGCACCAAGTTCCGCGAGACAAGGCACCCGCTGTACCGCGCGCTGCGGCGCGGGGCAACG  
 CGGAACGGTGGGTATGCGAGGTGCGCGTCCCGGCAAGCGCGGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCC  
 GAGATCGCAGCGCGCGCAACGATGCCGCAATGCTCGCCCTGGGCGGCGGCTCCGCCGCGCGCTCAACTTCTCGGA  
 CTCCGCGTGGTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGGCGGTGAGGCCGTCT  
 CGGATTTGACGCGACGGGAGGCTGCCGGTGGGTCCATCAGCGCCACGGTCGACGAGGAGGCCTCCTGTGGCGCTCCT  
 GCAGAATCGTCTGTGAGTCTGACGGTGCCGGCTCGTTCGAGACGTTCGAAACCTTCCGCCGATGGAGACTTGGCGGT  
 ACCGGTTCGGAATGGACATCGAAATGTTGAGGCTTGACTTCTTCCCGGAAATGGAGTTTGGCTCGTACTACGCGAGCC  
 TCGCGGAGGCGTGTCTATGGACCCGCGCGCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAATTCCGG  
 GAGGTGGCAACCGAGTTCCGCGCTGTGGAGCTAG

>HvCBFIIIc-10A (AK357794) (HarunaNi jo)

ATGACATGGGCCTTGAGGTCTCGAGCTCTCCCCCTCTCTTCCAACGAGAACGCGTCGGGACGGTCTGTCGACGGC  
 CAAGCGCCCCGGCGGGGCGCACCAAGTTCCGCGAGACAAGGCACCCGCTGTACCGCGCGCTGCGGCGCGGGGCAACG  
 CGGAACGGTGGGTATGCGAGGTGCGCGTCCCGGCAAGCGCGGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCC  
 GAGATCGCAGCGCGCGCAACGATGCCGCAATGCTCGCCCTGGGCGGCGGCTCCGCCGCGCGCTCAACTTCTCGGA  
 CTCCGCGTGGTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGGCGGTGAGGCCGTCT  
 CGGATTTGACGCGACGGGAGGCTGCCGGTGGGTCCATCAGCGCCACGGTCGACGAGGAGGCCTCCTGTGGCGCTCCT  
 GCAGAATCGTCTGTGAGTCTGACGGTGCCGGCTCGTTCGAGACGTTCGAAACCTTCCGCCGATGGAGACTTGGCGGT  
 ACCGGTTCGGAATGGACATCGAAATGTTGAGGCTTGACTTCTTCCCGGAAATGGAGTTTGGCTCGTACTACGCGAGCC  
 TCGCGGAGGCGTGTCTATGGACCCGCGCGCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAATTCCGG  
 GAGGTGGCAACCGAGTTCCGCGCTGTGGAGCTAG



TCGCGGAGGCGCTGCTGATGGACCCGCCCGCTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAATTC  
GGGAGGTGGCAACCGAGTTCGCGCTGTGGAGCTAG

>HvCBFIIIc-10A (AY785884) (88Ab536)

ATGGACATGGGCCTTGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGAACGCGTCGGGACGGTCTGTCGACGGT  
CAAGCGCCCGCGGGGCGCACCAAGTTCGCGGAGACAAGGCACCCGCTGTACCGCGCGGTGCGGCGCCGGGGCAACG  
CCGAACGGTGGGTATGCGAGGTGCGCGTCCCGGGCAAGCGCGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCC  
GAGATCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGGCGCGTCCGCCGCGCGCTCAACTTCGGGA  
CTCCGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGCGGTTCGAGGCCGTCTG  
CGGACTTGACAGCGACGGGAGGCTGCCGGTGGGTCCATCAGCGCCACGGTCGACGAGGAGGCCTCCTGTGGCGCTCCT  
GCAGAATCGTCTGCTGAGTCTGACGGTGCCGGCTCGTCGGAGACGTGAAACCTTCCGCGGATGGAGACTTGGCGGT  
ACCGGTCGGAATGGACATCGAAATGTTACGGCTTGACTTCTTCCCGAAATGGAGTTGGCTCGTACTACGCGAGCC  
TCGCGGAGGCGCTGCTCATGGACCCGCCCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAATTCGGG  
GAGGTGGCAACCGAGTTCGCGCTGTGGAGCTAG

>HvCBFIIIc-10A (AY785881) (Morex)

ATGGACATGGGCCTTGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGAACGCGTCGGGACGGTCTGTCGACGGC  
CAAGCGCCCGCGGGGCGCACCAAGTTCGCGGAGACAAGGCACCCGCTGTACCGCGCGGTGCGGCGCCGGGGCAACG  
CCGAACGGTGGGTCTGCGAGGTTCGCGTCCCGGGCAAGCGCGCGCTCGGCTCTGGCTCGGGACGTACGCCACGGCC  
GAGATCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGGCGCGTCCGCCGCGCGCTCAACTTCGGGA  
CTCCGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGCGGTTCGAGGCCGTCTG  
CGGATTGACAGCGACGGGAGGCTGCCGGTGGGTCCATCAGCGCCACGGTCGACGAGGAGGCCTCCTGTGGCGCTCCT  
GCAGAATCGTCTGCTGAGTCTGACGGTGCCGGCTCGTCGGAGACGTGAAACCTTCCGCGGATGGAGACTTGGCGGT  
ACCGGTCGGAATGGACATCGAAATGTTACGGCTTGACTTCTTCCCGAAATGGAGTTGGCTCGTACTACGCGAGCC  
TCGCGGAGGCGCTGCTCATGGACCCGCCCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAATTCGGG  
GAGGTGGCAACCGAGTTCGCGCTGTGGAGCTAG

>HvCBFIIIc-10A (DQ445248) (Tremois)

ATGGACATTGGCCTTGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGATCGCGTCGGGACGGTCTGTCGACGGC  
CAAGCGCCCGCGGGGCGCACCAAGTTCGCGGAGACAAGGCACCCGCTGTACCGCGCGGTGCGGCGCCGGGGCAACG  
CCGAACGGTGGGTCTGCGAGGTTCGCGTCCCGGGCAAGCGCGCGCTCGGCTCTGGCTCGGGACGTACGCCACGGCC  
GAGATCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGGCGCGTCCGCCGCGCGCTCAACTTCGGGA  
CTCCGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGCGGTTCGAGGCCGTCTG  
CGGATTGACAGCGACGGGAGGCTGCCGGTGGGTCCATCAGCGCCACGGTCGACGAGGAGGCCTCCTGTGGCGCTCCT  
GCAGAATCGTCTGCTGAGTCTGACGGTGCCGGCTCGTCGGAGACGTGAAACCTTCCGCGGATGGAGACTTGGCGGT  
ACCGGTCGGAATGGACATCGAAATGTTACGGCTTGACTTCTTCCCGAAATGGAGTTGGCTCGTACTACGCGAGCC  
TCGCGGAGGCGCTGCTCATGGACCCGCCCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAATTCGGG  
GAAGGTGGCAACCGAGTTCGCGCTGTGGAGCTAGTCTTGGACTCTGCTTTGTTTGA

>HvCBFIIIc-10B (AY785888) (Optia)

ATGGACATGGGCGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGAACGCGTCGGGACGGTCTGTCGACGGCAA  
GCGCCCGCGGGGCGCACTAAGTTCGCGGAGACAAGGCACCCGCTGTACCGCGCGGTGCGGCGCCGGGGCAACGCG  
AACGGTGGGTATGCGAGGTGCGCGTCCCGGGCAAGCGCGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCCGAG  
ATCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGTCCGCTCCGCCGCGCGGCTCAACTTCCCGGACTC  
CGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGCGGTTCGAGGCCGTCTGCGG  
ACTTGACAGCGACGGGAGGCCGCCGGTGGGTCCATCACCGCCACAGTCGGGAGGAGGCCTCCCGTGGCGCTCCTGCA  
GAATCGTCTGCTGAGTCTGAGGATGCCGGTTCGTCGGAGACGTGAAACCTTCCGCCGTTGGAGACTTGGCGGTACC  
GGTCGGAATGGACGTGAAATGTTACGGCTTGACTTCTTCCCGAAATGGAGTTGGCTCGTACTACGCGAGCCTCG  
CGGAGGCGCTGCTCATGGACCCGCCCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAGTTCGGCGAG  
GTGGCAACTGAGTTCGCACTCTGGAGCTAG

>HvCBFIIIc-10B (DQ445248) (Tremois)



ATGGACATGGGCGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGAACGCGTCGGGACGGTCGTCGACGGC  
 CAAGCGCCCGGGCGGCGCACTAAGTTCCGCGAGACAAGGCACCCGCTGTACCGCGCGTGC GGCGCGCGGBCAACCG  
 CGAACCGGTGGGTATGCGAGGTGCGCGTCCCGGCCAAGCGCGGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCG  
 GAGATCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGTCCGCTCCGCGCGCGGGCTCAACTTCCCGGA  
 CTCGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGGCGGTGAGGCCGTGCG  
 CGGACTTGACGCGACGGGAGGCCGCCGTGGGTCCATCACCGCCACAGTCGGGGAGGAGGCTCCCGTGGCGCTCCT  
 GCAGAATCGTCGTCTGAGTCTGAGGATGCCGGTTCGTGCGAGACGTGAAACCTTCCGCCGGTGGAGACTTGGCGGT  
 ACCGGTCGGAATGGACGTGAAATGTTAGGCTTGACTTCTTCCCGAAATGGAGTTGGCTCGTACTACGCGAGCC  
 TCGCGGAGGCGCTGCTGATGGACCCGCCGTGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAGTTTCGGC  
 GAGGTGGCAACTGAGTTCGCACTCTGGAGCTAG

>HvCBFIIIc-10B (AY785887) (88Ab536)

ATGGACATGGGCGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGAACGCGTCGGGACGCTCGTCGACGGCAA  
 CGCGCCCGGGCGGCGCACTAAGTTCCGCGAGACAAGGCACCCGCTGTACCGCGCGTGC GGCGCGCGGGCAACGCCG  
 AACCGTGGGTATGCGAGGTGCGCGTCCCGGCCAAGCGCGGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCCGAG  
 AGCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGTCCGCTCCGCGCGCGGCTCAACTTCCCGGACTC  
 CGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGGCGGTGAGGCCGTGCGGG  
 ACTTGACGCGACGGGAGGCCGCCGTGGGTCCATCACCGCCACAGTCGGGGAGGAGGCTCCCGTGGCGCTCCTGCA  
 GAATCGTCGTCTGAGTCTGAGGATGCCGGTTCGTGCGAGATGTCGAAACCTTCCGCCGGTGGAGACTTGGCGGTACC  
 GGTCGGAATGGACGTGAAATGTTAGGCTTGACTTCTTCCCGAAATGGAGTTTGGCTCGTACTACGCGAGCCTCG  
 CGGAGGCGCTGCTGATGGACCCGCCCGGTGGCGAACAGCACGGGCGCGTACTGGGATAACGGAGAGTTTCGGCGAG  
 GTGGCAACTGAGTTCGCACTCTGGAGCTAG

>HvCBFIIIc-10B (DQ445241) (Nure) (DQ445232) (Dicktoo) (AY785886) (Strider)

ATGGACATGGGCGAGGTCTCGAGCTCTCCCCCTCCTCTTCCAACGAGAACGCGTCGGGACGGTCGTCGACGGCAA  
 CGCGCCCGGGCGGCGCACTAAGTTCCGCGAGACAAGGCACCCGCTGTACCGCGCGTGC GGCGCGCGGGCAACGCCG  
 AACCGTGGGTATGCGAGGTGCGCGTCCCGGCCAAGCGCGGCGCTCGGCTCTGGCTCGGGACTTACGCCACGGCCGAG  
 ATCGCAGCGCGCGGAACGATGCCGCAATGCTCGCCCTGGGCGTCCGCTCCGCGCGCGGCTCAACTTCCCGGACTC  
 CGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGCGCGGGCGGTGAGGCCGTGCGGG  
 ACTTGACGCGACGGGAGGCCCTCCCGTGGCGCTCCTGCAGAATCGTCGTCTGAGTCTGAGGATGCCGGTTCGTGCGAG  
 ACGTCGAAACCTTCCGCCGGTGGAGACTTGGCGGTACCGGTGCGGAATGGACGTGAAATGTTAGGCTTGACTTCTT  
 CCCGAAATGGAGTTTGGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTGATGGACCCGCCCGGTGGCGAACA  
 GCACGGGCGCGTACTGGGATAACGGAGAGTTTCGGCGAGGTGCAACTGAGTTCGCACTCTGGAGCTAG

>HvCBFIIIc-13 (DQ095158) (Dicktoo)

ATGGGCATGGATCTTTCAGCTCTCCCCGTCTCATCGGTGTCATCTCGCCAGAGCACGCATCGGGGCCGGCGAA  
 CGCGCCCGGTGGGGCGCACTAAGTTCCGGGAGACCCGGCACCCGCTGTACCGACGCTGCGTCGAGGGGCAACGCCG  
 GACGGTGGGTCTGCGAGGTGCGCGTCCCGAGCAAGCGTGGCGCTCGGCTCTGGCTCGGGACGTACCTACCCGCCGGG  
 GCCGCGCGCGCGCAACGACGCGCGCATGCTCGCACTCGGCGCGCGCTCCGCCAGGCGCCTCAACTTCGCGGACTC  
 CGCGTGGCTGCTCGCCGTGCCGTTCGCGCTCTCCGACCTCGCCGACGTCCGGCGCACCGGCCTCCAGGCCGTAGCGA  
 ATTTCCAGCGACGGGAGGCTGCCAGTGGCCTGATAACCAGGACCGTCGCCGACAGTGTGATTATCGGAGACGTGCG  
 GAACCTTCCGCCGATGGAGATTTGAGTTGCCGGTTGCGATGGACAGCGATATGTTACGGCTTGACTTTTGGCCGGA  
 AATGGACCTAGGCTCGTACTACGTGAGCCTCGCGGAGGCGCTGCTCATGGACCCGCCGTGACGGCCACCATCATCG  
 ACGCGACCGGGACAACGGCGACGGCGCAAAGGTTTTCTTTTGGGAGAAGACTCTGTATAGCAAAGATTTTGAA  
 GCTGTTTTCAATCCATCACTCAATCCAGAAAAAGGAAATATACCTTCACGGCTTCATGTGCCATGA

>HvCBFIIId-12 (DQ095157) (Dicktoo)

ATGGACACGGTCCCGAGCGTAAGTGAAGTTCGCCGGCATCCCGCCGTCTCCTCGAGCAGGGGATGCCGTGCTC  
 TCTGCGTCGCCGACGCCGAAGCGCCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGCTGTTCACGGCG  
 TGGCGCGCGCGGCGAGCAACGGCAGGTGGGTGTGCGAGGTGCGCGTGGCGGGGAAGCGCGCGGAGCGGCTCTGGCTC  
 GGCACGCACGTACCGCGGAGGCGGGCGCGCGCACGACGCCCGCATGCTCGCGCTGTACGGGCGCACACCTGC  
 TGGCGCGCTCAACTTCCCGGACTCGGCGTGGCTGCTCGCCGTGCCCTCCTCCTCTCCGACATGGCGGACGTCCGGC







CGGAGGGGCTTCTCCTGGAGCCGCCGAACCGCCGCCGCCGCCGCTGCTGGGATACCGATGGCGGTGGAGCA  
GACGCCGCGCTCTGGAGCTACTAA

>HvCBFIIId-15B (EU593539) (Tremois)

ATGGACATGGCCGGCTCCGATCAGCAGCGGTGCTCCCCTTCCTCGCCGTCGTTGTCTCTCATCTGAAGCGCCCCG  
CGGGCGCACCAAGTTCAAGGAGACGCGTCACCCG GTGTACCGCGCGGTGCGGCGCCGCGGCAGCGCCGGCCGGTGGG  
TGTGCGAGGTGCGCGTCCCCGGCAAGCGAGGCGAGCGGCTCTGGCTCGGGACGCACCTTACCGCCGAGGCGGCCGCG  
CGTGCGCACGACGCCGCCATGCTCTGCCTGCTCGACCGCCGCCGCCCGTGTCTCAACTTTGCC GACTCTGTCTGGCT  
CCTCGCCGTGCCGTCCGCGCTCTCCGATCTCGCTGATGTCCGGCGCGCGGCTCTCAGCGCCGTGCGGATTTCAGC  
GCCGGGAGGCCGCAAGCGGCGCCGCCACGAGGGCCAGGCGGCGGCCGCCCTCATTGACGAGGGAACCTGTAGCCAA  
TCCGCGCAGTCTGTCATGGAATAACCGGCTCGTCTTCGACGTCGTATCCCTACCTTACGCCGACGGAATGCTTGA  
GGTGCCGGCCACACTGGGCAGCAACATGTTGAGCTGGACATGTCCGGGAAATGGACCTGGACACATACTATGCGT  
ACTTCGCGGAGGGGCTTCTCCTGGAGCCGCCGAACCGCCGCCGCCGCCGCGCTGCTGGGATATCGATGGCGGTGGA  
GCAGACGCCGCGCTCTGGAGCTACTAA

>HvCBFIIId-15A (EU593535) (Nure)

ATGGCCGGCTCCGATCAGCAGCGGTGCTCCCCTTCCTCGCCGTCGTTGTCTCTCATCTGAAGCGCCCCGCCGGCG  
CACCAAGTTCAAGGAGACGCGTCACCCG GTGTACCGCGCGGTGCGGCGCCGCGGCAGCGCCGGCCGGTGGGTGTGCG  
AGGTGCGCGTCCCCGGCAAGCGAGGCGAGCGGCTCTGGCTCGGGACGCACCTTACCGCCGAGGCGGCCGCCGCGTGGC  
CACGACGCCGCCATGCTCTGCCTGCTCGACCGCCGCCGCCCGTGTCTCAACTTTGCC GACTCTGTCTGGCTCTCGC  
CGTGCCGTCCGCGCTCTCCGATCTCGCTGATGTCCGGCGCGCGGCTCTCAGCGCCGTGCGGATTTCAGCGCCGGG  
AGGCCGCAAGCGGCGCCGCCACGAGGGCCAGGCGGCGGCCGCCCTCATTGACGAGGGAACCTGTAGCCAATCCGCG  
CAGTCGTCCATGGAATAACCGGCTCGTCTGACGTCGTATCCCTACCTTACGCCGACGGAATGCTTGAGGTGCC  
GACCACACTGGGCAGCAACATGTTGAGCTGGACATGTCCGGGAAATGGACCTGGACACATACTATGCGTACTTCG  
CGGAGGGGCTTCTCCTGGAGCCGCCGAACCGCCGCCGCCGCCGCGCTGCTGGGATACCGATGGCGGTGGAGCAGAC  
GCCGCGCTCTGGAGCTACTAA

>HvCBFIIId-15A (AK375451) (HarunaNijo)

ATGGACATGGCCGGCTCCGATCAGCAGCGGTGCTCCCCTTCCTCGCCGTCGTTGTCTCTCATCTGAAGCGCCCCG  
CGGGCGCACCAAGTTCAAGGAGACGCGTCACCCG GTGTACCGCGCGGTGCGGCGCCGCGGCAGCGCCGGCCGGTGGG  
TGTGCGAGGTGCGCGTCCCCGGCAAGCGAGGCGAGCGGCTCTGGCTCGGGACGCACCTTACCGCCGAGGCGGCCGCA  
CGTGCGCACGACGCCGCCATGCTCTGCCTGCTCGACCGCC CGCCCCGTGTCTCAACTTCGCCGATTCCGCTCTGGCT  
CCTCGCGTGCCGTCTCGCGCTCTCCGATCTCGCTGACGTCCGGCGCGCGGCTCTCAGCGCCGTGCGGATTTCAGC  
GCCGGGAGGCCGCAAGCGGCGCCGCCACGAGGGCCAGGCGGCGGCCGCCCTCATTGACGAGGGAACCTGTAGCCAA  
TCCGCGCAGTCTGTCATGGAATAACCGGCTCGTCTTCGACGTCGTATCCCTACCTTACGCCGACGGAATGCTTGA  
GGTGCCGGCCACACTGGGCAGCAACATGTTGAGCTGGACATGTCCGGGAAATGGACCTGGACACATACTATGCGT  
ACTTCGCGGAGGGGCTTCTCCTGGAGCCGCCGAACCGCCGCCGCCGCCGCGCTGCTGGGATATCGATGGCGGTGGA  
GCAGACGCCGCGCTCTGGAGCTACTAA

>HvCBFIIId-16 (EU593542) (Tremois)

ATGGACATGACCGGTGCGACCAAGTGAAGCTCCTCTTCCTCGCCGTCATCGACCTCTCGCACCGGAAGCGCCC  
CGCCGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG GTGTACCGCGGTGTCGCGCGCCGGGGCAACGCGGCCGCT  
GGGTGTGCGAGGTGCGGGTCCCTGGGCAGCGCGGCGAGCGGCTTTGGCTCGGCACGTACCTTACCGCCGACGCGGCC  
GCACCGCGCGACGATGCCGCCATGATCGGCTGCTCGGCCACTCAGCCGCGTGCCCTCAACTTCCC GACTCCGCGTG  
GCTCCTGGCCGTGCCACCCGCGCTCTCCGACCTCGCGGCCGTCCGGCGCGGCCCTCGCCGCCGTAGCGGACTTCC  
AGCGGCGGCATGCCGCAACGGCGCAGCCACTGTCCCTGCCGATGAGGACACCTCCAGCGCGGACAATGCGGGCGGC  
TCGTGCGCGACGTCTCAGCTTCGGCCGAGGGGACGTTGCAAGTGCCATCCGCGCTGGGCAACGACATGTTGAGCT  
GGACTTGTCTGGGAGATGGACCTGGGCACGTACTACGCCGACCTCGCGGAGGGGATGCTCCTGGAGCCGCCGCCG  
CGCTGGACAGCGGGGCGTGTGGGATGCCGAGACGGCGGAGCTGACTACGGGCTATGGAGCTACTGA

>HvCBFIIId-16 (EU593536) (Nure)



ATGGACATGACCGGGTCGGACCAGCAATGGAGCTCCTCTTCTCGCCGTCATCGACCTCCTCGCACCCGAAGCG  
 CCCCGCCGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGGTGTGCGGCGCCGGGGCAACGCCGGCG  
 CTTGGGTGCGCGAGGTGCGGGTCCCTGGGCAGGCGCGGAGCGGCTTTGGCTCGGCACGTACCTCAACCGCGACCGC  
 GCGGCACGCGCGCACGATGCGCGCATGATCGGCCTGCTCGGCCACTCAGCCGCGTGCTCAACTTCCCGGACTCCGC  
 GTGGCTCTGGCCGTGCCACCCGCGCTCTCCGACCTCGCGGCGGTCCGGCGCGCGGCCCTCGCCGCGGTAGCGGACT  
 TCCAGCGGGCGCATGCCGGAACGGCGCAGCCACCGTCCCTGCCGATGAGGACACCTCCAGCGCGGACAATGCGGGC  
 GGCTCGTCGGCGACGTCTCAGCCTTCGGCCGAGGGGACGTTCAAGTGCCATCCGCGCTGTGCAGCGACATGTTCGA  
 GCTGGACTTGTCTGGGGAGATGGACCTGGGCACGTACTACGCCGACCTCGCGGAGGGGATGCTCCTGGAGCCGCCG  
 CGTCGCTGGACAGCGGGGCGTGCTGGGATGCCGAGACGGCGGAGCTGACTACGGCTATGGAGCTACTGA

>HvCBFIVa-2A (AY785841) (Dicktoo)

ATGGACACAGTTGCCGCTGGCCGCAAGTTGAGGAGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
 GACGGTTTGGTCGGAGCCGCGGAAGCGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGCTGTACCGCG  
 GCGTGCGACGCGGTGGCAAGGTGCGGCAGTGGGTGTGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
 TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGCGCGCGCGCACGACTCCGCGCGCTCGCCCTCTCCGGCCATGA  
 TGGCTGCTCAACTTCGCCGACTCCGCTGGCGGATGATGCCCGTCCACGCGACTGGGTCTGTTAGGCTCGCCCCCG  
 CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCAGAGTGTTCAGGGGCGAGCACCAGCCGACGCTGCACGGCC  
 GAGGAGAGCAGACCCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGACTGGATCGGCGGCATGGACGC  
 CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGCGCGCGGAGGGTGGCGGGAGGACGACG  
 CGAACACGACGACGGCTTCAACACGTCCGCGTCGCTGTGGAGCTACTAG

>HvCBFIVa-2A\_ (DQ480160) (Nure)

ATGGATACAGTTGCCGCTGGCCGCAAGTTGAGGAGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
 GACGGTTTGGTCGGAGCCGCGGAAGCGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGCTGTACCGCG  
 GCGTGCGACGCGGTGGCAAGGTGCGGCAGTGGGTGTGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
 TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGCGCGCGCGCACGACTCCGCGCGCTCGCCCTCTCCGGCCATGA  
 TGGCTGCTCAACTTCGCCGACTCCGCTGGCGGATGATGCCCGTCCACGCGACTGGGTCTGTTAGGCTCGCCCCCG  
 CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCAGAGTGTTCAGGGGCGAGCACCAGCCGACGCTGCACGGCC  
 GAGGAGAGCAGACCCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGACTGGATCGGCGGCATGGACGC  
 CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGCGCGCGGAGGGTGGCGGGAGGACGACG  
 CGAACACGACGACGGCTTCAACACGTCCGCGTCGCTGTGGAGCTACTAG

>HvCBFIVa-2A\_ (AK374123) (HarunaNijo)

ATGGACACAGTTGCCGCTGGCCGCAAGTTGAGGAGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
 GACGGTTTGGTCGGAGCCGCGGAAGCGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGCTGTACCGCG  
 GCGTGCGACGCGGTGGCAAGGTGCGGCAGTGGGTGTGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
 TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGCGCGCGCGCACGACTCCGCGCGCTCGCCCTCTCCGGCCATGA  
 TGGCTGCTCAACTTCGCCGACTCCGCTGGCGGATGATGCCCGTCCACGCGACTGGGTCTATTTAGGCTCGCCCCCG  
 CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCAGAGTGTTCAGGGGCGAGCACCAGCCGACGCTGCACGGCC  
 GAGGAGAGCAGACCCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGACTGGATCGGCGGCATGGACGC  
 CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGCGCGCGGAGGGTGGCGGGAGGACGACG  
 CGAACACGACGACGGCTTCAACACGTCCACGTGCTGTGGAGCTACTAG

>HvCBFIVa-2A\_ (GU461587) (GoldenPromise)

ATGGACACAGTTGCCGCTGGCCGCAATTTGAGGGGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
 GACGGTTTGGTCGGAGCCGCGGAAGCGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGCTGTACCGCG  
 GCGTGCGACGCGGTGGCAAGGTGCGGCAGTGGGTGTGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
 TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGCGCGCGCGCACGACTCCGCGCGCTCGCCCTCTCCGGCCATGA  
 TGGCTGCTCAACTTCGCCGACTCCGCTGGCGGATGATGCCCGTCCACGCGACTGGGTCTGTTAGGCTCGCCCCCG  
 CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCAGAGTGTTCAGGGGCGAGCACCAGCCGACGCTGCACGGCC  
 GAGGAGAGCAGACCCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGACTGGATCGGCGGCATGGACGC  
 CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGCGCGCGGAGGGTGGCGGGAGGACGACG  
 CGAACACGACGACGGCTTCAACACGTCCACGTGCTGTGGAGCTACTAG



CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGGCCCGGAGGGTGGCGGGAGGACG  
ACGGCGAACACGACGACGGCTTCAACACGTCCACGTCGCTGTGGAGCTACTTAG

>HvCBFIVa-2A\_ (DQ445249) (Tremois)

ATGGACACAGTTGCCGCTGGCCGCAATTTGAGGGGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTTTGGTCGGAGCCGCCGAAGCGGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGGTGTACCGTG  
CGTGCGACGCCGTGGCAAGGTTCGGGCAGTGGGTGTGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGGCGCGCGCACGACTCCGCCCGCGCTCGCCCTCTCCGGCCATGA  
TGGTGCCTCAACTTCGCCGACTCCGCCTGGCGGATGATGCCCGTCCACGCGACTGGGTCGTTCAGGCTCGCCCCCG  
CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCGAGGTGTTCCAGGGGCAGCACCCAGCCGACGCGTGCACGGCC  
GAGGAGAGCACGACCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGCACTGGATCGGCGGCATGGACGC  
CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGGCCCGGAGGGTGGCGGGAGGACGACG  
CGAACACGACGACGGCTTCAACACGTCCGCGTCGCTGTGGAGCTACTTAG

>HvCBFIVa-2A\_ (AY785843) (88Ab536)

ATGGACACAGTTGCCGCTGGCCGCAATTTGAGGAGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTTTGGTCGGAGCCGCCGAAGCGGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGGTGTACCGCG  
CGTGCGACGCCGTGGCAAGGTTCGGGCAGTGGGTGTGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGGCGCGCGCACGACTCCGCCCGCGCTCGCCCTCTCCGGCCATGA  
TGGTGCCTCAACTTCGCCGACTCCGCCTGGCGGATGATGCCCGTCCACGCGACTGGGTCGTTCAGGCTCGCCCCCG  
CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCGAGGTGTTCCAGGGGCAGCACCCAGCCGACGCGTGCACGGCC  
GAGGAGAGCACGACCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGCACTGGATCGGCGGCATGGACGC  
CGAGCACTGGATCGGCGGCATGGACGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCCG  
CCGCCGGAGGGTGGCGGGAGGACGACGGCGAACACGACGACGGCTTCAACACGTCCGCGTCGCTGTGGAGCTACTTAG

>HvCBFIVa-2B\_ (DQ097684) (Dicktoo)

ATGGACACAGTTGCCGCTGGCCGCAATTTGAGGGGCAAGACTACATGACGGTGTGGCCGGAGGAGCAGGAGTACCG  
GACGGTTTGGTCGGAGCCGCCGAAGCGGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGGTGTACCGCG  
CGTGCGACGCCGTGGCAAGGTTCGGGCAGTGGGTATGCGAGCTGCGCGTCCCGTAAGCCGGGGTTACTCCAGGCTC  
TGGCTCGGCACCTTCGCCAACCCCGAGATGGCGGGCGCGCGCACGACTCCGCCCGCGCTCGCCCTCTCCGGCCATGA  
TGGTGCCTCAACTTCGCCGACTCCGCCTGGCGGATGATGCCCGTCCACGCGACTGGGTCGTTCAGGCTCGCCCCCG  
CGCAAGAGATCAAGGACGCCGTGCGCGTCGCCCTCGAGGTGTTCCAGGGGCAGCACCCAGCCGACGCGTGCACGGCC  
GAGGGGAGCACGACCCCATCACCTCAAGCGACCTATCGGGGCTGGACGACGAGCACTGGATCGGCGGCATGGACGC  
CGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGGCTGCCGGAGGGTGGCGGGAGGACGACG  
CGAACACGACGACGGCTTACGACGTCCACGTCGCTGTGGAGCTACTTAG

>HvCBFIVc-14 (DQ095159) (Dicktoo)

ATGGACGCCGACAGCGCCGCTCCCGTGTGATGGGCACAGGACGGTGTGGTCGGAGCCGCCGAAGCGGCCTGCCGG  
CCGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGCAGCGCCGGGGCCCCGAGGCCGTTGGGTGT  
TCAAGTGCAGCTGCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGC  
TGGCACGACGCCCGCTTCTCGCGTCTCTGGCCGCGCCGCTGTCTGAACCTTCGCCGACTCTGCTTGGCGGATGCT  
CCCCCTCTTCGCCGGCCGTTTCAGCACCGCTAAGGAGATCAAGGATGCCGTGCGCGTCCCGTCTGGCGTTCCAAA  
GACAGCACCCGGTCCGTCACCGCGCCAATGTCCCCGGCACGACAGCCGTTGACGAGAAGGAAGTCGATGGCTCG  
CCGGCGCCGAGCGCTCTGTTTCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCAT  
CTACTCGGAGGGCATGTTTCATAGAGTCGCCGGATACCAGACCGTGGCGGAAGACCTCGAGCTCGGTGGCGTCCAG  
CACCGCCATGGAGCTACTTGTTCGACTAA

>HvCBFIVc-14 (DQ151545) (Morex)

ATGGACGCCGACAGCGCCGCTCCCGTGTGATGGGCACAGGACGGTGTGGTCGGAGCCGCCGAAGCGGCCTGCCGG  
CCGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGCAGCGCCGGGGCCCCGAGGCCGTTGGGTGT  
TCAAGTGCAGCTGCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGC  
TGGCACGACGCCCGCTTCTCGCGTCTCTGGCCGCGCCGCTGTCTGAACCTTCGCCGACTCTGCTTGGCGGATGCT



CCCCCTCCTCGCCGGCCCGTTTACGACCCGCTAAGGAGATCAAGGATGCCGTGCGCGTCGCCGTCTGGCGTTCC  
 AAAGACAGACCCCGTTCGCGTCCACGGCGCAATGTCCCCGGCACGGACAGCCGTTGACGAGAAGGAAGTCGATGGC  
 TCGCCGGCGCCGAGCGCTCTGTTTCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATC  
 ATGCTACTCGGAGGGCATGTTTCATAGAGTCGCCGGATACCAGACCGTGGCGGGAAGACCTCGAGCTCGGTGGCGTCC  
 AGACACCGCCATGGAGCTACTTGTTCGACTTAA

>HvCBFIVc-14 (EU650231) (Tremois)

ATGGACGCCGACAGACGACGCTCCCCGTGTGATGGGCACAGGACGGTGTGGTTCGGAGCCGCCGAAGCGGCCTGCCGG  
CCGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGCGGCGCCGGGGCCCCGAGGCCGGTGGGTGT  
 CCGAGGTGCGCGTGCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGC  
 GCGCACGACGCCGCCGTTCTCGCGCTCTCTGGCCGCGCGCGCTGTCTGAACCTTCGCCGACTCTGCTTGGCGGATGCT  
 CCCCCCTCCTCGCCGGCCCGTTTACGACCCGCTAAGGAGATCAAGGATGCCGTGCGCGTCGCCGTCTGGCGTTCCAAA  
 GACAGACCCCGTTCGCGTCCACGGCGCAATGTCCCCGGCACGGACAGCCGTTGACGAGAAGGAAGTCGATGGCTCG  
 CCGGCGCCGAGCACTCTGTTTCATGTCCAGCGAGCTGTTAAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCATG  
 CTACTCGGAGGGCATGTTTCATAGAGTCGCCGGATACCAGACCGTGGCGGGAAGACCTCGAGCTCGGTGGCGTCCAGA  
 CACCGCCATGGAGCTACTTGTTCGACTTAA

>HvCBFIVd-4A/B (AY785849) (Dicktoo)

ATGGACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGAAGCAGCAGGGGCACAGGACGGTGTGCTCGGAGCCGCC  
GAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGCGGCGCCGGGGCGGG  
 TCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGCTCGGGACCTTCACCAACCCC  
 GAGATGGCCCGCGCGCCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCTGCCTCAACTTCGCCGACTC  
CGCGTGGCGGATGCGGCCAGTGCTCGCGACCACCGGGTCGTTCCGCTTCAGCAGCACGCGGGAGATCAAGCTTGCCG  
 TCGCCGTAGCCGTGCTCGCGTTCCAGCAGCAGCAGATTATTCTTCCAGTAGCGTGTCCATCGCCGGAGGCGCCCGCC  
 AGCCCCGAGCGCCGCTCTGTTTTACATCTCGTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGAGGCAT  
 GGACGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCCGCGGACGAAAGAGCGAGGCCAGAGA  
 ACCGCGAGCACAGCGGCGTCGAGACACCAATACCACTATGGAGCTATTTGTTTCGACTGCTAA

>HvCBFIVd-4A/B (AK359618) (HarunaNi jo)

ATGGACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGAAGCAGCAGGGGCACAGGACGGTGTGCTCGGAGCCGCC  
GAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGCGGCGCCGGGGCGGG  
 TCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGCTCGGGACCTTCACCAACCCC  
 GAGATGGCCCGCGCGCCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCTGCCTCAACTTCGCCGACTC  
CGCGTGGCGGATGCGGCCAGTGCTCGCGACCACCGGGTCGTTCCGCTTCAGCAGCACGCGGGAGATCAAGCTTGCCG  
 TCGCCGTAGCCGTGCTCGCGTTCCAGCAGCAGCAGATTATTCTTCCAGTAGCGTGTCCATCGCCGGAGGCGCCCGCC  
 AGCCCCGAGCGCCTCTGTTTTACATCTCGTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGAGGCAT  
 GGACGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCCGCGGACGAAAGAGCGAGGCCAGAGA  
 ACCGCGAGCACAGCGGCGTCGAGACACCAATACCACTATGGAGCTATTTGTTTCGACTGCTAA

>HvCBFIVd-4A/B (EU593538) (Tremois)

ATGGACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGAAGCAGCAGGGGCACAGGACGGTGTGCTCGGAGCCGCC  
GAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGCGGCGCCGGGGCGGG  
 TCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGCTCGGGACCTTCACCAACCCC  
 GAGATGGCCCGCGCGCCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCTGCCTCAACTTCGCCGACTC  
CGCGTGGCGGATGCGGCCAGTGCTCGCGACCACCGGGTCGTTCCGCTTCAGCAGCACGCGGGAGATCAAGCTTGCCG  
 TCGCCGTAGCCGTGCTCGCGTTCCAGCAGCAGCAGATTATTCTTCCAGTAGCGTGTCCATCGCCGGAGGCGCCCGCC  
 AGCCCCGAGCGCCGCTCTGTTTTACATCTCGTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCAT  
 GGACGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCCGCGGACGAAAGAGCGAGGCCAGAGA  
 ACCGCGAGCACAGCGGCGTCGAGACACCAATACCACTATGGAGCTATTTGTTTCGACTGCTAA

>HvCBFIVd-4A/B (AY785850) (Dicktoo)



**ATG**GACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGAAGCAGCAGGGGCACAGGACGGTGTCTGTCGGAGCC  
 CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCTGTGTACCGCGCGGTGCGGCGCGCGGGCC  
 GGTCTGGGCACTGGGTGTGCGAGGTGCGCGTGGCGGGGATCAAGGGCTCCAGGCTCTGGCTCGGGACCTTCACCAAC  
 CCGAGATGGCCGCGCGCGCCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCCGA  
 CTCGCGTGGCGGATGCGGCCAGTGCTCGCGACCACCGGGTCGTTTGGCTTCAGCAGCACGCGGGAGATCAAGCTTG  
 CCGTCGCCGTAGCCGTCTGCGGTTCCAGCAGCAGCAGATTATTCTTCCAGTAGCGTGTCCATCGCCGGAGGCGCCC  
 GCCAGCCCGAGCGCGCTCTGTTTTACATCTCGTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGAGG  
 CATGGACGCCGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCAG  
 AGAACCGGAGCACAGCGGCGTCGAGACACCAATACCACTATGGAGCTATTTGTTGACTGCTAA

>HvCBFIVd-4A/B (AY785848) (Morex)

**ATG**GACGTCGCCGACATCGCCTCCCCGTCCGGCCAGCAGAAGCAGCAGGGGCACAGGACGGTGTCTGTCGGAGCCCGC  
 GAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCTGTGTACCGCGCGGTGCGGCGCGCGGGCCGGG  
 TCGGGCAGTGGGTGTGCGAGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTCGGGACCTTCACCAACCC  
 GAGATGGCCGCGCGTSCCCACGACGCGCGGTGCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCCGACTC  
 CGCGTGGCGGATGCGGCCAGTGCTCGCGACCACCGGGTCGTTTGGCTTCAGCAGCACGCGGGAGATCAAGCTTGCCG  
 TCGCCGTAGCCGTCTGCGGTTCCAGCAGCAGCAGATTATTCTTCCAGTAGCGTGTCCATCGCCGGAGGCGCCCGCC  
 AGCCCGAGCGCCGTCTGTTTTACATCTCGTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCAT  
 GGACGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCAGAGA  
 ACCCGGAGCACAGCGGCGTCGAGACACCAATACCACTATGGAGCTATTTGTTGACTGCTAA

>HvCBFIVd-4D (AY785852) (88Ab536)

GCCTCCCCGTCCGGCCAGCAGGAGCAGCAGGGGCACCGGACGGTGTCTGTCGGAGCCGCGAAGCGGCCCGCGGGGCG  
 CACCAAGTTCCACGAGACGCGCCACCCTGTGTACCGCGCGGTGCGCGCGCGCGGTGCGGGCAGTGGGTGTGCG  
 AGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACTTCAACACGGCCGAGATGGCCGCGCGCGG  
 CACGACGCGCGCTCGCTCGCGCTCTCCGGCCGCGCGCGCTGCCTCAACTTCGCCGACTCCGCGTGGCGGATGCGGCC  
 AGTGCTCGCGACCACCGGGTCGTTTCGGCTTCAGCAGCACGCGGGAGATCAAGCTTGCCGTGCGCGTAGCCGTCTGTCG  
 CGTTCCAGCAGCAGCAGATTATTCTTCCAGTAGCGTGTCCATCGCCGGAGGCGCCCGCCAGCCGAGCGCCGCTCTG  
 TTTTACATCTCGTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCCGGGTCTGACTA  
 CGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGAGACGCGGAGCAGACCGGCG  
 TCCAGACACCAATACCACTATGG

>HvCBFIVd-9 (AY785878) (Dicktoo)

**ATG**TGAATCCGATTAGACGGACGTCGCCGGCATCGCCTCCCCGTCCGGCCAGCAGGAGCAGCAGGGGCACCGGAC  
 GGTGTCTGTCGGAGCCGCGAAGCGGCCCGCGGGGCGCACCAAGTTCCACGAGACGCGCCACCCTGTGTACCGCGCG  
 TCGCGCGCGCGCGCGCGTGGGCACTGGGTGTGCGAGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTC  
 GGCACCTTCAACACGGCCGAGATGGCCGCGCGCGCACGACGCGCGCGCTCGCGCTCTCCGGCCGCGCGCGCTG  
 CTTCAACTTCGCCGACTCCGCGTGGCGGATGCTGCCCCTGCTCGCGGCCGGGTGCTTCGGCTTCGGCAGCGCGCGG  
 AGGTCAAGGCCGCGCTCGCGTTCGGCTCGTGGCGTTCAGCGGAGGAGATTATTCAGTCCCGTTCGCCGTCTGTCG  
 GCGCTCCAGAAGCAGCAGGTTCCGGTCTGTCGCCGTCTGACGCTCCAGCAGAAGCAGCAGAGTCCCGGTTCG  
 CGTTCGCCGTTCGGCGCTCCAGCAGCAGAGTCCCGTTCGCCGTTCGCCGTTCGGCGCTCCAGCAGTTCAGGTT  
 CGGTTCGCCGTTCGGCTCGTGGCGCTCCAGGAGCAGAGATTATTCTACAGTTCGGTGCCTGGCGCCGAGTTCTAC  
 ATGTCTTCAGGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTTCGGCGGAATGGACGCCGGGTCTACTACGCGAG  
 CTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGAGACGCGGAGCAGACCGGCGTCCAGA  
 CGCCGCTATGGAGTGTCTGTTGACTAA

>HvCBFIVd-9 (EU332037) (Kompoltikorai)

**ATG**TGAATCCGATTAGACGGACGTCGCCGGCATCGCCTCCCCGTCCGGCCAGCAGGAGCAGCAGGGGCACCGGAC  
 GGTGTCTGTCGGAGCCGCGAAGCGGCCCGCGGGGCGCACCAAGTTCCACGAGACGCGCCACCCTGTGTACCGCGCG  
 TCGCGCGCGCGCGCGCGTGGGCACTGGGTGTGCGAGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTC  
 GGCACCTTCAACACGGCCGAGATGGCCGCGCGCGCACGACGCGCGCGCTCGCGCTCTCCGGCCGCGCGCGCTG  
 CTTCAACTTCGCCGACTCCGCGTGGCGGATGCTGCCCCTGCTCGCGGCCGGGTGCTTCGGCTTCGGCAGCGCGCGG



AGGTCAAGGCCCGCTCGCCGTCGCCGTCGTGGCGTTCCAGCGGAGGCAGATTATTCCAGTCGCCGTCGCCGTC  
GTGGCGCTCCAGAAGCAGCAGGTTCCGGTCGCTGTTCGCCGTCGTGACGCTCCAGCAGAAGCAGCAGCAGGTTCCCGT  
CGCGCTCGCCGTCGTGGCGCTCCAGCAGCAGCAGGTTCCCGTCCCGTCCCGTCCCGTCCAGCAGCTGCAGG  
TTCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGGAGCAGCAGATTATTCTACAGTCGCTGCCTGGCGCCGAGTTT  
TACATGTCTTACGGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTTCGGCGGAATGGACGCCGGGTCGTACTACGC  
GAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGGAGCACGGCGAGCAGACCGGCGTCC  
AGACGCCGCTATGGAGCTGCTTGTTCGACTTAA

>HvCBFIVd-9 (EU332038) (OWB-D)

ATGTCGAATCCGATTTCAGACGGACGTCGCCGGCATCGCCTCCCCGTCCGGCCAGCAGGAGCAGCAGGGGACCGGAC  
GGTGTCTGTCGGAGCCGCCGAAGCGGCCCGCGGGGCGCACCAAGTTCCACGAGACGCGCCACCCTCTGTACCGGGT  
TGGCGCGCGCGCGCGCTCGGGCAGTGGGTGTTCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTC  
GACACCTTCAACACGGCCGAGATGGCCGCCCGCGCGCACGACGCCGCCGCGCTCGCGCTCTCCGGCCGCGCGCTG  
CTCAACTTCGCGGACTCCGCTTGGCGGATGGTGCCCGTGTCTCGCGCCGGGTCGTTCCGGCTTCGGCAGCGCGCGG  
AGGTCAAGGCCCGCTCGCCGTCGCCGTCGTGGCGTTCCAGCGGAGGCAGATTATTCCAGTCGCCGTCGCCGTCGTG  
GCGCTCCAGAAGCAGCAGGTTCCGGTCGCTGTTCGCCGTCGTGGCGCTCCAGCAGCAGCAGGTTCCGGTCCGCCGTCGC  
CGTCGTGGCGCTCCAGCAGCTGCAGTCCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGGAGCAGCAGATTATTCTAC  
CAGTCGCGTGCCTGGCGCCGAGTTTACATGTCTTACGGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTCCGGC  
GGAATGGACGCCGGGTCGTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCC  
GGAGCACGGCGAGCAGACCGGCTCCAGACGCCGCTATGGAGCTGCTTGTTCGACTTAA

>HvCBFIVd-9 (EU332039) (OWB-R)

ATGTCGAATCCGATTTCAGACGGACGTCGCCGGCATCGCCTCCCCGTCCGGCCAGCAGGAGCAGCAGGGGACCGGAC  
GGTGTCTGTCGGAGCCGCCGAAGCGGCCCGCGGGGCGCACCAAGTTCCACGAGACGCGCCACCCTCTGTACCGGGT  
TGGCGCGCGCGCGCGCTCGGGCAGTGGGTGTTCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTC  
GACACCTTCAACACGGCCGAGATGGCCGCCCGCGCGCACGACGCCGCCGCGCTCGCGCTCTCCGGCCGCGCGCTG  
CTCAACTTCGCGGACTCCGCTTGGCGGATGCTGCCCGTGTCTCGCGCCGGGTCGTTCCGGCTTCGGCAGCGCGCGG  
AGGTCAAGGCCCGCTCGCCGTCGCCGTCGTGGCGTTCCAGCGGAGGCAGATTATTCCAGTCGCCGTCGCCGTCGTG  
GCGCTCCAGAAGCAGCAGGTTCCGGTCGCTGTTCGCCGTCGTGACGCTCCAGCAGAAGCAGCAGCAGGTTCCCGGTCGC  
CGTCGCCGTCCTCGCGCTCCAGCAGCAGCAGGTTCCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGTGCAGGTTT  
CGGTGCCGTCGCCGTCGTGGCGCTCCAGGAGCAGCAGATTATTCTACAGTCGCTGCCTGGCGCCGAGTTTCTAC  
ATGTCTTCAGGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTCCGGCGGAATGGACGCCGGGTCGTACTACGCGAG  
CTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGGAGCACGGCGAGCAGACCGGCGTCCAGA  
CGCCGCTATGGAGCTGCTTGTTCGACTTAA

>HvCBFIVd-9 (EU332036) (Scarlett)

ATGTCGAATCCGATTTCAGACGGACGTCGCCGGCATCGCCTCCCCGTCCGGCCAGCAGGAGCAGCAGGGGACCGGAC  
GGTGTCTGTCGGAGCCGCCGAAGCGGCCCGCGGGGCGCACCAAGTTCCACGAGACGCGCCACCCTCTGTACCGGGT  
TGGCGCGCGCGCGCGCTCGGGCAGTGGGTGTTCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTC  
GACACCTTCAACACGGCCGAGATGGCCGCCCGCGCGCACGACGCCGCCGCGCTCGCGCTCTCCGGCCGCGCGCTG  
CTCAACTTCGCGGACTCCGCTTGGCGGATGCTGCCCGTGTCTCGCGCCGGGTCGTTCCGGCTTCGACAGCGCGCGG  
AGGTCAAGGCCCGCTCGCCGTCGCCGTCGTGGCGTTCCAGCGGAGGCAGATTATTCCAGTCGCCGTCGCCGTCGTG  
GCGCTCCAGAAGCAGCAGGTTCCGGTCGCTGTTCGCCGTCGTGACGCTCCAGCAGAAGCAGCAGCAGGTTCCCGGTCGC  
CGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGGTTCCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGTGCAGGTTT  
CGGTGCCGTCGCCGTCGTGGCGCTCCAGGAGCAGCAGATTATTCTACAGTCGCTGCCTGGCGCCGAGTTTCTAC  
ATGTCTTCAGGCGACCTGCTGGAGCTCGACGAGGAGCAGTGGTTCCGGCGGAATGGACGCCGGGTCGTACTACGCGAG  
CTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGGAGCACGGCGAGCAGACCGGCGTCCAGA  
CGCCGCTATGGAGCTGCTTGTTCGACTTAA

>ScCBFI-11 (EU194240) (Puma)

CAGCGCTATTTCATCGTCGGGAACGAGTTTCCGGCCGCCGGCGACCGGGAGGACGGCCCGTACATGACGGTGTCTGT  
CGCGCCGCCGAAGCGCGGGCTGGGAGGACCAAGGTCAGGGAGACGAGGCACCCGCTGTACAGGCGGGTGGCGAGC

GGAACCCCGGCGGGTGGTCTGCGAGGTGCGCGAGCCGACCGGGAAGCAGAGGCTATGGCTCGGCACCTTCGA  
 ACCGCGGAGATGGCGCGCGCGCACGACGTGCGCGCCCTGGCGCTCCGCGGCGCGCAGCGTGCCTCAACTTCG  
 TGAATCGCCGCGCACGCTCCGGGTGCCGCCGACAGGGGGCCGGCCACGACGAGATACGCCGGGCGCGGTAGAAGCG  
 GCCGAGCTGTTCCGCGCGCGCGCGGAGCCCAATGCAGCTGCCGAGGCGCCGGCTGCTTCGCCGGTGGCATCGGG  
 GAACGCGGAGCTCGTCGAAAGCTCTCCTTACTGCCTCATGGATGGACTGGAATTCGAAATGCAGGGCTATCTTGACA  
 TGGCGCACGGCATGCTCATCGAGCCACCGCCAATGGCGGGGCCGTCTACGTGGATCGAGGAGGACTACGACTGCGAG  
 G

>ScCBFI-11 (HQ730766) (Lo152)

TACAGCGGGAGCGACTACTCGTCGTCGGGAGCGCAGTCTCCGTTGGCCGGCGGCCGGGAGGAGGGCTCGTACATGAC  
 CGTGTGTCGTCGGCGCCGCCAAGCGGGCGGGCGGGAGGACCAAGGTCAGGGAGACGAGGCACCCGCTGTACAAAGGGG  
 TCGCAGCAGGAACCCCGGACGGTGGGTCTGCGAGGTGCGCGAGCCGACCGGGAAGCAGAGGCTATGGCTCGGCACC  
 TCGACACCGCGCGAGATGGCGGGCGCGCGCACGACGTGCGCGCCCTGGCGCTCCGCGGCCGCGCAGCGTGCCTCAA  
 TTCCCGGACTCGCCGCGCACGCTCCGGGTGCCGCCGAGGGGGCCGGCCACGACGAGATACGCCGGGCGCGGTAG  
 AAGCGGCCGAGCTGTTCCGCGCGCGCGCGGAGCCCAATGCAGCTGCCGAGGCGCCGGCTGCTTCGCCGGTGGCA  
 TCGGGGAACGCGGAGCTCGTCGCAAGCTCTCCTTACTGCCTCATGGATGGACTGGAATTCGAAATGCAGGGCTATCT  
 TGACATGGCGCACGGCATGCTCATCGAGCCACCGCCAATGGCGGGGTCTACGTGGATCGAGGAGGACTACGACT  
 GCGA

>ScCBFII-5 (EU194241) (Puma)

CGGTGTGGCCTACTATGGCAGCACCGCCGGCGGTAGACGACAACGGCCAGGGCGGGTACGCGACGGTGACCTCGG  
 CGCCGCCCAAGCGGGCGGGCGGGCGGACCAAGTTCCGGGAGACGCGGCACCCGCCGTACCGGGCGGTGCGCAGGCGC  
 GCGCGCGCGGGGCGGTGGGTGTGCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCGGCACCTTCGCCAG  
 TCCGGAGGCCGCGGCCCGCGCCACGACGTGGCCGCGCTGGCCCTGCGCGGCCGCGCCGCTGCCTCAACTTCGCCG  
 ACTCGGCCACGCTGCTCGCCGTGACCCCGCCACGCTGCGCACCCCGACGACATCCGCGCCGCGGCCATCGCGCTC  
 GCCGAGACCGCCTGCCCGCGCGCGCCGCGCGTCTCTGTCGTCGTTGGCCGCCGGGGCGCCCCGATGATGGCGAT  
 GATGCAGGAGTCCGCGCGCGCGCCGTACGACAGCTACGCCATGTACGGCGGCTTGGCGGACCTAGACCACCATTCCT  
 ACTACCACTACGACGGATGAGCTGTCGGCGGGCGGCGACTGCCAGAGTATCTCCACATGAACGGAGCCGACGAA  
 GACGGCAGCTACGGCGCAGGAGACG

>ScCBFIIIa-6 (EU194242) (Puma)

AATGAGCGGGGAGTCCGGCTCGCCGTGACGCGGCGAGAATTCTACTCGCCCTCCACGTGCGCGGAGCATCAGCAGG  
 CGAGGCCGACTGTGTGACGTCGCGCGCCCAGGAAGCGGCCGCGGGGCGGACCAAGTTAGGGAGACGCGGCACCCG  
 TTGTACCGCGGCGTGCAGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCTGGCAGGCGCGGAG  
 TAGGCTCTGGCTCGGCACCTTCGACACCGCGGAGGCGCGCGCGCGCAACGACGCCACCATGATCGCGCTCAACG  
 TCGGGGGCGCGCGTGCCTCAACTTTGCCGACTCGGCCGAGCTGCTCGCCGTGCCGGCGGCCTCTCTTACCGCAGC  
 CTGGACGAGGTCCGCCACGCCGTGTCGAGGCGGTCGAGGACTTCTTGGCGGCCAGGCGATCGCCGAGGACGACGC  
 GCTCTCGGGACCTCGTCGTCGCGCCCTCTCTCTACCGACGACGAGTCGTCTCTTCCCCGGCGGAGGCCGAGG  
 CCGAGGACTCGCCGTTGAGCTGGAGCTCCTGAGCGACATGGGCTGGGACCTGTACTACGCGAGCTGGCGCAGGCG  
 ATGCTCATGGCGCGCCCTCCGCGGCGGCGGCTCGGCGATTGCGGCGAGGCTACCTCGCCGACGTGCCAC

>ScCBFIIIa-6 (HQ730764) (Lo152)

CTGTGTGACGTCGCGCGCCGGGAAGCGCGCGGGGCGGACCAAGTTAGGGAGACGCGGCACCCGGTGTACCGC  
 GGCGTGGCGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGACAGGCGGGGAGCAGGCTCTG  
 GCTCGGCACCTTCGACACCGCGGAGGCCGCGCGCGCCAAACGACGCCACCATGATCGCGCTCAACGCCGGGGCG  
 CCGCGTGCCTCAACTTTGCCGACTCGGCCGAGCTGCTCGCCGTGCCGGCGGCCTCTCTTACCGCAGCTGGACGAG  
 GTCCGCCACGCCGTGTCGAGGCGGTCGAGGACTTCTTGGCGGCCAGGCGATCGCCGAGGACGACGCGCTCTCGGG  
 CACCTCGTCGTCGCGCCCTCTCTCTACCGACGACGAGTCGTCTCTTCCCCGGCGGAGGCCGAGGCCGAGGACT  
 CGCCGTTGAGCTGGACGTCCTGAGCGACATGG

>ScCBFIIIc-3A (EU194244) (Puma)



CTCGAGCTCCTCCCTCTCCTCCTCGTTCGGTGTCTGCTCCTCGCCCGACACGCATCGGAGCGA **GTGAAGCGCCCG**  
**CGGGGCGCACCAAGTTCCGGGAGACGCGGACCCG** **GTGTACCGGCGGTGGGCGCGGTGGCAACACCAACCGTGG**  
**GTGTGGAGGTTCGGGCTCCGCGCAAGCGTGGCGCGGAGGCTCTCGCTTCGCAAGTACGCGACCGCTGAGATCCGCGC**  
**GGCGCGCAACGATCCGCGCCTCTGCTCTGGGCGCGGTTCGGCGCGGTGCCTCAAGTTCCGA** **GA** **ACTCCGCGTGGC**  
**TG**CTCGCCGTGCGGCTCCTCGCTCTCGACCTCGGCGACGTCCGGCGCGCGCGGTTCGAGGCCGTTCGCGGATTTCCAG  
ACACGAGAGGTTGCCAACGGCTCCCTCGCGGCCACCGTCACCGAAGAGGCTCCTGTGGCGCTCCTGAAGAATCGTC  
GTCTGAGTCTGACAGTCCGGTTTCGTTCGGAGACGTTCGGAACCTTCAGCCGATAGAGAGTTGGAGGTGCCGCTCGCAG  
TGGACACCGATATGTTCCGGGCTTGACTTGTTCGCGAAATGGACCTGTGCTCGTACTACGCGAGCCTCGCGGAGGCG  
CTGCTCGTGGACCCGCCGCGCACCGGTGACCACCAACGACGTGTACTGGGACAATGGCGACGGCGGAGCT **TGATG**

>**ScCBFIIIc-3B (EU194245) (Puma)**

CTCGAGCTCCTCCCATCCTCCTCGTTCGGTGTCTGCTCCTCGCCCGAGCACGCGGCGGGGCGAG **CG**TCGCTGCCC **AAGC**  
**GC** **CCCCCGGGGGCGCACCAAGTTCCGGGAGACGCGGACCCG** **GTGTACCGGCGGTGGGCGCGGTGGCAACACCAACCGTGG**  
**TCGTGGGTCTGGGAGGTTCGGGCTCCGCGCAAGCGTGGCGCGAGGCTCTGGCTTCGGGACGTACGCCACGGCTGAGAT**  
**TCGCGCGCAACCGAAGCGCGCGCATGCTCGCGCTCGGCGCGCGCTTCGCGCGGTGCCTCAACTTCGCA** **GA** **ACTCCG**  
**CGTGGCTG**CTCGCCGTGCGGCTCCTCGCTCTCGACCTCGGCGACGTCCGGCGCGCGCGGTTCGAGGCCGTTCGCGGAT  
TTCCAGACACGAGAGGTTGCCAACGGCTCCCTCGCGGCCACCGTCACCGAAGAGGCTCCTGTGGCGCTCCTGAAGA  
ATCGTCTGCTGAGTCTGACAGTCCCGGTTTCGTTCGGAGACGTTCGGAACCTTCAGCCGATAGAGAGTTGGAGGTGCCGCG  
TCGCACTGGACACCGATATGTTCCGGGCTTGACTTATTCCGCGAAATGGACCTGTGCTCGTACTACGCGAGCCTCGCG  
GAGGCGCTGCTCGTGGACCCGCCGCGCACCGGTGACCACCAACGACGTGTACTGGGACAATGGCGACGGCGGAGCT **TGA**  
**TG**

>**ScCBFIIIe-10 (EU194243) (Puma)**

TCCTCAGCTCCTCTTCCAACGAGAACGCGTCGACG **GCCAAGCGTCCGGCGGGGCGCACCAAGTTCCGCGAGACAAGG**  
**CACCCG** **GTGTACCGCGCGGTTCGGCGCAGGGGCAACGCGCGAGCGGTGGGTCTGCGAGGTTCGCGCTCCCGGCAAGCG**  
**CGGCGCGAGGCTCTGGCTTCGGGACGTACGCCACGGCGGAGATCGCAGCGCGCGCAACGATGCCGCCATGCTCGCCC**  
**TGGGCGGCGCTTCGCGCGCGCGCTCAACTTCCCG** **GA** **ACTCCGCGTGGCTG**CTCGCCGTGCGGCTTCGCGCACTCCGAT  
CTGGCCGACGTCCGGCGCGCGCGGTTCGAGGCCGTTCGCGGATTTGCAGCGACGGGAGGCTTCGCGCGGGTCCATCAC  
CGCCACCGCGCGCGAGGACGCTCCTGTGGCGCTCCTGCAGAATCGTCTGAGTCTGACGATGCCGGTTTCGTTCGG  
AGACGTGAAACCTTCTGCCGATGGAGAGTTCGCGGTTCGCGGCGGAATGGACATCGAAATGCTCAGGCTTGACTTG  
TTCCCGGAAATGGACTTGGGCTCGTACTACGTGAGCCTTCGCGAGGCGCTGCTCATGGACCCGCTTCGGTGGCGAC  
CGCACCGGCGCGTACTGGGATAACGGAGAGTTCGCGCGAGGGGGCA **ACTG**

>**ScCBFIIId-12 (EU194246) (Puma)**

CAACTGGAACCTCGCCGCATCCCCCGCTCGTCTGCTCGAGCAGGGGATGCCGTCTGCTCGCTTCGCTCGCCGACG **CCGA**  
**AGCGCCCCCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG** **GTGTTCACGGCGTTCGCGCGCGGGGAAGCAAC**  
**CGCGCGTGGGTGTGCGAGGTTCGCGGTTCGCGGGAAGCGCGCGGAGCGGCTCTGGCTTCGGGACGCACGTACCGCGCGA**  
**GGCAGCGCGCGCGCACACGACGCGCGCATGCTTCGCGCTGTACGGCGCGCAACCCCTCCATGCGCTCAACTTCCCGG**  
**ACTCGGCGTGGCTG**CTCGCCGTGCCCTCCTCCTCTCCGACCTGGCGGACGTCCGGCGCGCGCGCATCGGGGAGTTC  
GTGGACTTCTGCGCGGCGAGGAACGGGCGCGCGCGCTGGCGCGCGCGATGAGGTGGCCCCCATCGACGGGATTCG  
CTACGCTGCGTTCGGCGCGGACAATGCCAGCTCGTTCGGCGCGGAGTCTCACTCGCAGATGCTTGGCGCAATGCGG  
AGTTTGAGTTGCCGACGCAATTGTGCCACGACATGTTGACCTCCACAGTTCGCGGAAATGGACGCGGGCACGTAC  
TACGCGGACCTCGCGCAGGGGCTGCTCCTGGAGCGCGCGCGCGCGCTCCAGCAGGGCGAGCTCGGAGCACGGAGA  
**TGATG**

>**ScCBFIIId-12 (HQ730767) (Lo152)**

**ATG**GACACGGGCCCCGAGCGCAACTGGAACCTCGCCGCATCCCCCGCTCGTCTGCTCGAGCAGGGGATGGCGTCTGTC  
GCTTGCCTCGCCGACG **CCGAAGCGCCCCCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCG** **GTGTTCACGGCG**  
**TGCGCGCGCGGGGAAGCAACGCGCGGTGGGTGTGCGAGGTTCGCGGTTCGCGGGAAGCGCGCGGAGCGGCTCTGGCTG**  
**GGCAGCGACGTACCGCGGAGGCGCTTCGCGCGCGCATGACCGACCCATGATACCGGTGTACGGCGCGCACCGCGG**  
**CTC** **CGCGCTCAACTTCCCG** **GA** **CTCGGCGTGGCTG**CTCGCCGTGCCCTCCTCCTCCTCGACCTGGCGGATGTCCGGC  
GCGCGCCATCGGGGAGTCTGTGGAATTCCTGCGCGGCGAGGAGACGGGCGCGCGGTGTCTGCCNNNNNNNNNNNN



TCTGCAGCGGAGCTCTCCTTCTCTCCGTCGTCTCTCTCGCACCTGAAGCGCCCGGCCGGGCGCACCAAGTTCAAGG  
AGACGCGCCACCCGCTGTACCGCGCGGTGCGCGCGCGCGGTAGCGCGCGCGGTGGGTGTGCGAGGTTCCCGTCCCG  
TGAAGCGCGCGGAGCGGCTGTGGCTCGGCACGCACCTCACTGCCAGGCGGCGCGCGGAGCGCACGACGCCGCCAT  
CTCGGCGCTGATCGGCCCCCTCCACCGCGTGTCTCAACTTCGCGGACTCCGCCTGGCTCTCGCGGTGCCGTCCGCGC  
TCTCCGACTTTGCAGACGTCCGGCGCGCGGCTCTCAGCGTGTGCGAGACTTCCAGCGCCGGGAGACCGCCAGCGGC  
GCCGCCACCGTTCCCGTCGATGACGGAACTGTAGCCAAATCCGCGCAGTCGTCCATGGAATAACCGGCTCGTCTGTG  
GACGTCGTCGTCCCTATCTCTTGGCGACGGAATGTTTGAGGTGCCAGCACACTGGGCTGCGACATGTTTCGAGCTGG  
ATATGTCGGGGAAATGGACCTGGACATACTATGCTACTTCGCGGAGGGGCTTCTCCTTGAGCCGCCGCAATCG  
TTCGCTCGTCGGCGCGTGTCTGGGACACTGAAGCGCGCGGAGCAGACG

TGCACCAAGTCTCAGGAGACGCGCCACCCGGTGTACCGCGGCGTTCGGCGGCGCGGTAGCGCCGCGCGGTGGGTGTG  
 TCTCTTCGCGTCCCGCGCAAGCGCGGCGAGCGGCTGTGGCTCGGCACGCACCTCACTGCCGAGGCGGCGCGCGAG  
 TCCCGCACGCGCCCATGCTCGGCCGTGATCGGCCCTCCACCCCGTGTCTCAACTTCGCCGACTCCGCCTGGCTCTCTC  
 CGCGTGCCGTCCGCGCTCTCCGACTTTGCAGACGTCCGGCGCGCGGCTCTCAGCGCTGTTCGACAGACTTCCAGCGCCG  
 TGAGACCGCCAGCGGCGCGGCCACCGTTCCCGTCGATGACGGAACTGTAGCCAATCCGCGCAGTCGTCCATGGAAA  
 TACCGGCTCGCTCGTGACGCTCGTCTGCTCCCTATCTTCTGGCGACGGAATGTTGAGGTGCCGACCACACTGGGCTGC  
 ACATGTTCTGAGCTGGATATGTCCGGGGGAAATGGACCTGG

TCGATCAGCTCCCTTCTCATCGAGCTCCGGGCACGAGCACGGGAAGCGGTTCTGTGTGGTCGCCGGCGCGCA  
TGGGCCCGCGGGSCGACCAAGTTCAAGGAGACGCGCCACCCG GTGTACCGCGGCGTGC GGCGCCGGGGCAGCGCG  
TGGGTGTGCGAGGTGCGCGTTCGCCGGAAGCGCGCGAGCGGCTCTGGCTCGGGACCTACGTCGCCCGCAGA  
TGC CGCGCGCCTCAGCAGCGCGCGGATGCTCGCCCTGCTTGGACGCTCCCTTCCGCGCGCGCGGTGCCTCAACT  
GACTCCGCGTGGCTGCTCGTCATGCCCCGAGGCTCTCCGACCTGGCCGACGTCCGGCGTGC GGCCATCGAG  
CCG TAGCGGCGTTCTGTGCGCTGGAGGCGCGCGCGTGGTCCCGACGCTCTACGACGCGACCTCCCCGTGTACCT  
CCCTGCGCCGTGGGCAATGCCAGCAAGTGTTCCAGTGCCGACTTCTCCCCGTGGGCAGCGACATGTTGAGC  
CGACATGTCCGGGGAAGTGATCTGGATGCTGCTACTACGTCAGGACTTTGCCAGGGGATGCTTGTGAGCCGCGCCC  
TCCCGCGTACTGGGAGAACGGAGAATGCGGCGACGGCGGAGCGGCCT

TATGCGTTCGGGCGGGGGCGGAACAAGTTGCAGGAGACACGCCACCCAAGTGATACCGCGCGGTGCGCCCGCGTGGCCG  
 TAAATCTTTTGGTGTTCAGCTGCGGTGCGGGCGGAAGTCCGAGTTACTCCAGGATCTGCGTGGCAAGCTTCC  
 TCAATCTTCAATATGTCGGCGGCTCCAGGAATCGCCCGCGCTCGCGCTTTCCGGCCCGGATGCGTGCCCTCAACTTC  
 GCGTAATCTCGGCCCTGGCGGATGATGCCCGTCCACGCGCGCGGTCTGTTCAAGTTGGCCGCCGCTCAGGAGATCAAGGA  
 TCTCGCCGTCGCCCTCGAGGCGTTCCAGGAGCAGCAGCGCCCTGCCGACGTGTCAACGGCGACGTCTGTCCACGG  
 TAGGAGAGCGCGCTCTCCATCATCCCCAGCGACCTGTCTGGGGCTCGACAGCGAGCACTGGATCGGCGGCATGGAG  
 TCCGATCATACTACGCGAGCTTGGCGCAGGCGATGTCTATGGAGCCGCCACCGACGGAGCGTGGCGGGAGGACCG  
 TACACGACGACGAGTTCGACACGTCGCTGTGGAGCTACTAG

**CCTAAGCGTGGGCGGGGCGGAACAAGTTGCAGGAGACACGCCACCAGTGTAACGGCGGGGTGGCCGCCGTGGCCG**



CCGAGGAGAGCGCTCTCCATCATCCCCAGCGACCTGTCTGGGGCTCGACAGCGAGCACTGGATCGGCGGCATG  
GAGGCCGGGTCTACTACGCGAGCTTGGCGCAGGCGATGCTCATGGAGCCGCCGACCGAGCGGTGGCGGGAGGA  
CCGCGAACACGACAACGGATTTCGACACGTCGCTGTGGAGTTACTAG

>ScCBFIVa-2 (HQ730763)

ATGGACACCGTTTGGCGCCCGGCCGATTTTGAAGGGCAAGAGTACATGACGGTGTGGCCAGAGGAGGAGTACCGGAC  
CGTGTGGTCCGAGCCGCGAAGCGGGCGGGCCGGAACAAGTTGCAAGAGACGCGCCACCCGCTGTACCGGGCG  
TGCGCCCGCGTGGCGGGGAAGGGCAGTGGGTGTGGGTGTGAGCTGCGAGTGGCCGGCCGCGGAAGCCGGGTTTAC  
TCCAGGATCTGGCTCGGCACCTTCGCGGACCCGAGCATGGCGGGCGCGCGCACGACTCGGGCCGGCTCGCGCTCTC  
TGGCCCGGATGCTGGCTCAACTTCGCGGACTCTGCTGGCGGATGATGCCGGTGCACGCGGCCGGGTCTTCAAGT  
TGGCCGCGCGCAGGAGATCAAGGACGCCGTGCGCGTTCAGGCGTTCCAGGAGCAGCAGTGCCTTGGCGAC  
CGGTCAAAGCG

>ScCBFIVb-20 (AF370728) (clone79)

ATGGACGCGCGCAGCGCGGCTCCCCCGTTTTGGGCACAGGACGGTGTGCTCGGAGCCGCCAAGAGGCCGGCAGG  
GCGGACCAAGTTTAAAGGAGACCCGCCACCCGCTGTACCGCGCGTGGCGCGCGGGGTGGCTCGGGCAGTGGGTGT  
TCGAGGTGCGCGTGGCGGGCGCGCAAGGTACAGGCTCTGGCTCGGCACATTACACCCGCGGAGATGGCGGGCGCGC  
GCGCACGACTCCGCGGTGCTCGCGCTCTCGACCGCGCGCTTGCTCAACTTCGCGGACTCCGCTGGCGGATGCT  
GCCCGTCTCGCGGACGGCTCGTCCCGCTTACGACGCGCGCGGAAATCAAGGACGCCGTGCGCGTTCGCGTCTGTG  
AGTTCCAGCGGCAGCGCCCTTCGTGTCCACGTGCGAGACGGCGCAGCGGAGAAAGGACGTCCAAGGTCGCGGAGG  
CCGAGCGAGCTGTCCACGTCCAGCGACTTGTGGACGAGCACTGGTTTAGCGGCATGGACGCCGGCTCTTACTACGC  
GAGCTTGGCGCAGGGATGCTCATGGAGCCCGCGCGCCAGAGCGTGGAGCGAGGATGGCGGCGAATACAGCGGCG  
TCCACACGCCGCTTTGGAACTAG

>ScCBFIVc-14 (HQ730768) (Lo152)

TGGCGACGGGGCCCGCAGGCGCGGTGGGTGTGCGAGGTGCGCGTGTCTGGGATGAGCGGTTCCAGGCTCTGGCTCGG  
CACCTTCACCACCGCGGAGATGGCAGCGCGCGCGCACGACCGCGCGTTCTCGCGCTCTCCGGCCGCGCGCTTGCC  
TCAACTTCGCGGACTCTGCGCTGGCGGATGCTCCCGTCTCTCGCGGCGCGCTTCAGCACCGCGAAGGAGATCAAGGAT  
GCCGTGCGCGTTCGAGTCTTGGCGTTCCAAAGGAAGCACCCGGTTCGCGTCCACGACGCCATCGTCCCTGCGCGGAC  
AACCGATGACGAGAAGGAAATCGGTGTCTCGCGGCGCGGAGCGCCCTCTCATGTCTAGCGAGCTGTTGAATGAGC  
ACTGGTTTGGCGGCATGGATGCCGGATCGTGTACTCGGAGGCGATGTTACGGAGTCGCTGGACACCAGACCCCTGG  
CGGAAGACATCGAGTCCGTGGCTCGAGACACCGCCATGGAGCTGCTTGTTCGACTTAA

>ScCBFIVd-9 (AF370729) (clone80)

ATGGACGTGCGCGACATCGCCTCCCGTCTGGCCAGCAGCAGCAGGGGACCGGACCGTGTCTGTCGGAGCCGCCGAA  
GCGCCCCGCGGGGAGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGGCGCGCGTGGCGCGCTCG  
GCGAGTGGGTGTGCGAGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAG  
ATGGCGGCGCGCGCGCACGACGCTGCCGTGCTCGCGCTCTCCGCGCGCGCGCGCTGCCTCAACTTCGCGGACTCCGC  
CTGGCGGATGCTGCCCGTGTCTCGCGGCCGGCTCCTTCGGCTTTGATAGCGCGCGGAGGTCAAGGCCCGCGTTCGCG  
TCGCGCTCGTTCGCTTCCAGCGGAAACAGATTATTCCAGTCGCGCTCGTGTCTGCTCTCCAGAAGCAGCAGGTT  
CCGCTCGCGTGGCGTCTGGCGCTCCAGCAGAGGACAGTTCCGGTCACCGTCGCGTCTGGCGCTCCAGAAGCT  
GCAGGTTCCGGTCCCGTCCCGTCTGGCGCTCCAGAAGAAGCAGATTATTCTTCCAGCCGCGTGTCTGGCGCGCG  
AGTTTTACATGTCTTCCGGCGACCTGTTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCCGGGTCTGTAC  
TACGCCAGCTTGGCGCAGGGATGCTCGTGGCGCGCGCGGACGACAGAGCGAGGCCGAGAACGGCGAGCAGAGCGG  
CGTCCAGACTCCGCTATGGAGCTGCTTGTTCGACTTAA

>ScCBFIVd-9 (AF370730) (clone81)

ATGGACGTGCGCGACATCGCCTCCCGTCTGGCCAGCAGGAGCAGGGGACCGGACCGTGTCTGTCGGAGCCGCCGAA  
GCGCCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGGCGCGCGTGGCGCGCTCG  
GCGAGTGGGTGTGCGAGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAG  
ATGGCGGCGCGCGCGCACGACGCGTCTCGCGCTCTCTCGCGCGCGCGCGCTGCCTCAACTTCGCGGACTCCGC  
CTGGCGGATGCTGCCCGTGTCTCGCGGCCGGTCTTTCGGCTTCGGCAGCCCGCGGAGATCAAGGACGCCGTTCGCG  
CGTCCAGACTCCGCTATGGAGCTGCTTGTTCGACTTAA



[illegible]

TGCCCTCAACTTCGCC
GACTCCGCCTGGCGG
ATGCTGCCCGTGCTCGN
GGGCGGCTCCTTCGGCTTTGATAGCACGCG
GGAGGTCAAGGCCCGCTCGCCGTGCGCGTTCCAGCGGAAGCAGATTATTCCAGTCGCCGTGCGCCGTG
TGGCGCTCCAGCAGCAGAGGTTCCAGTCGGCGTGCGCGTCTGTTGGCGCTCCAGCAGAAGCAGGTTCCAGTCGCCGTG
GCCGTCGTGGCGCTCCAGCAGAAGCAGGTTCCGGTCGCCGTGCGCGCTCGTGGCGCTCCAGAAGCAGCAGATTATTCT
TCCAGCCGATGCCTGGCGCGGAGTTCTACATGTCTTCCGGCAGCTCTTGGAGCTCGACGAGGAGCAGTGGTTTG
GGCGGCATGGACGCCGGGGTCGTACTACGCCAGCTTGGCGCAGGGGATGCTCG

**Figure S3: Liste des séquences codantes de cultivars Norstar, Chinese Spring et Manitou (de *T. aestivum*) démontrant des polymorphismes spécifiques.** Les gènes CBF et le numéro du clone sont en surlignage rose et décrits dans le Tableau 3. Les acides nucléiques mutés sont en gris. Le codon d'initiation (ATG) et codon d'arrêt sont en rouge (s'ils sont présents). Les amorces utilisées sont en bleue. Toutes les séquences ont été traduites et analysées par *BLAST*.

>TaCBFI-1.2\_S06(manque\_301b\_pour\_ATG)

CGCTGCGCGGCCCGCCCGCTGCTCAACTTCGCGGACTCGCTCGGCGGCTCCGGGTCCCGCCCGTGGGTCTCT  
CCTGATGAAATACGGCGGGCGGGCGGTAGAGGCGACTGAGGCATTCTGCCGGCACCCGACCAGAGCAATGCC  
CGAGGAGGCAGCCGTTGCACCATCGGAGCAGTTTCGCCGTGATCCCTACTATGGGATGGACGATGGGATG  
GGATGCAGGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCCTCCGATGGCAGGCCCGTCGGCTCT  
GGTGACGGCGACGATGACGGCGAGGTCAGCTTGTGGAGCTACTGATTACGCGCATTGAGTTATGCCCTTGT  
GCGAAGAAGTTCTAATGCATGTGCAGCTTTATATAGCAATCCAGTACAGGAAGTGTACTGCATGGTTGCCCT  
TTTCTCTTTTCTGAACTTCTCTGGGTAATGTATGTTTGCTCTCTGATGAACTGCGATCAGAAGAAGCAGAA  
ATCTTGAGTTTGGTAAATGGTGACGCTTGTCCGGTTATGGATTGTGAGTGTTCGAACTT

>TaCBFI-1.2-S07(manque\_146b\_pour\_ATG)

GGAGACGCGCACCCCGCTCTACAGGGCGTCCGCGGAGGAACCCCGGGAGGTGGGTCTGCGAGGTGCGG  
ACACAAAGCAGAGGATATGGCTCGGCACGTTTCGAGACCGCAGAGATGGCGGCGCGCGCACGACGTGCT  
GCGCTGCGCGGCCGCGCCGCTGCCTCAACTTCGCGGACTCGCTCGGCGGCTCCGGGTCCCGCCCGTGGGT  
CCCTGATGAAATACGGCGGGCGGGCGGTAGAGGCGGCTGAGGCATTCTGCCGGCACCCGACCAGAGCAATG  
CCGAGGAGGCAGCCGTTGCACCATCGGAGCAGTTTCGCCGTGATCCCTACTATGGGATGGACGATGGGAT  
GGGATGCAGGGCTACCTCGACATGGCGCAGGGGATGCTCATTGCCCTCCTCCGATGGCAGGCCCGTCGCGCACTG  
CGGTGACGGCGACGATGACGGCGAGGTCAGCTTGTGGAGCTACTGATTACGCGCATTGAGTTATGCCCTTGTGGG  
CGGAAGAAGTTCTAATGCATGTGCAGCTTTATATAGCAATCCAGTACAGGAAGTGTACTGCATGGTTGCCGTTCCA  
CTTCTCTTTTCTGAACTTCTCTGGGTAATGTATGTTTGCTCTCTGATGAACTGCGATCAGAAGAAGCAGAA  
GATCTTGAGTTTGGTAAATGGTGACGCTTGTCCGGTTATGGATTGTGAGTGTTCGAACTCCG

>TaCBFII-5.1b-M08 (manque\_101b\_pour\_ATG)

GTTCGCGGAGACGCGGCACCCGGTGTACCGGGCGTGCGCCGGCGCGGCGCGGCCGGCGGTGGGTCTGCGAGGTGC  
GCGAGCCCAACAAGTCCCGCATCTGGCTCGGCACCTTCGCTCCCCGAGGCCGCGGCCCGCGCCACGACGTT  
GCCGCGCTGGCCCTGCGCGGCCGCGCCGCGTGCCTCAACTTCGCCGACTCGGCGGCCCTGCTCGCCGTCGACCCG  
CAGCTCCGACCCCGACGACATCCGCGCGGCGGCCATCGCGCTCGCCGAGACTGCCTGCCCCGCGCGCGGCC  
CGTCGTCGGCCGTGGCCGCCGTGGCGTCCGCGCCGGCGCCCCGATGATGACGATGATGTACGAGTCCGCGCGGT  
CACTACGACGACTACCCGATGCAGTACGGGTACGGCGGCATTGGGACTTGGACCAGGATTCTACTACTAC  
TCTACCGCCGCGCGGCGGCGACTGGCAGAGCGGCTCGCACATGGACGGAGCCGACGACGACTGCAACGGCAGCGGT  
TCTACCGCGCGCGGCGAGGTCCCGCTCTGGAGCTACTGATCGAGCCGCTCGATTGAGCTAGTCTCTAGCCCGATTGGC  
TCTCT

>TaCBFII-5.3\_M14 (manque\_251b\_pour\_ATG)

CGCCAGCCCCGAGGCCGCGGCCCGCGCCACGAGCTCGCCGCGCTCGCGCTCCGGGGCCGCGGCCCGCTGCT  
TCGCCGACTCGGCCACGCTGCTCGCCGTCGACCCCGCCACGCTCCGACCCCCGACGACATCCGCGCCGCCCGCT  
GCGCTCGCCGAGACCGGTGCCCGCTGCGCCCGCTCCTCGTCGTCGTTGGCCGCCCGGTGGCGTCCGCGCC  
GCCCCGATGACGATGATGCAGTTTGACGACTACGCGATGCAGTACGGCGGCATTGGGACTTGGACCAGCATT  
ACTACTACGACGGTTGAGCGCCGCGGTGGCGACTGGCCGAGCGGCTCGCACATGGACGGAGCCGACGACGACT  
AACGGCAGCGGTGGCTACGGCGCCGGCGAGGTGCGGCTCTGGAGCTACTGATCGAGCTGCTCGATTGAGCCATT  
CAGCCCGGTACGCTTT

>TaCBFIIa-6.1\_S29 (manque\_73b\_pour\_ATG)



CCACGTCGCGGGAGCATCAGCAGGCGAAGCAGGCGGGCTGGACGTCGGCGCCGGCGAAGCGGCCGGCGGGGGCGG  
 ACCAAGTTTACGGGAGACGCGGCACCCGGTATACCGCGCGCTGCGGCGCAGGGGCAATGCCGGGCGGTGGGTGTGCGA  
 GGTGCGCGTGGCCGGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCCA  
 ACACGCGCCCATGATCGCGCTCTCCGCCGGGGCGCGGGCTGCCCTCAACTTCGCCGACTCCGCCGAGCTGCTCGCC  
 GTGCCGGCGGCTCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTCTGGAGGCCGTCGAGGACTTCTCGCG  
 GCGGGAGGCGATCGCCGAGGAGACGCGCTCTCGGGCACCTCGTCGTCCGCGCCCTCTCCCTCACCGACGACGAGT  
 CGTCTCTTCCCCGCGGAGGACTCGCCGTTTCGAGCTGGATGTCTGAGCGACATGGGCTGGGACCTGTACTACGCG  
 AGCCTGGCGCAGGCGATGCTGATGGCGCGCCCTTCTTCCATGGCCGCGGCGCTCGGCGATTACGGCGAGGTCGATGT  
 GCCACTCTGTAGCTACCAGAGCTAGCTAGCTAGTTTCGCGCCAGTTCGAATATTTTAC

>TaCBFIIIIa-6.1\_M21

GCCAAGGAACAGAGCATCGAAGGAA GAAGAAG ATG TGTCCGATCAAGAAGGAGATGAGCGGGGAGTCCGGCTCGCCG  
 TGCAGCGCGGAGAACTTCTACTCGCCCTCCACGTCGCGGGAGCATCAGCAGGCGAAGCAGGCGGCGTGGACGTCCGC  
 GCCGCGAAGCGGCCGGCGGGGCGGACCAAGTTTCAGGGAGACGCGGCACCCGGTATACCGCGGCGTGCAGCGCAGGG  
 GCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGAC  
 ACCGCCGAGGCCGCCGCGCGCCAACGACGCGCCCATGATCGCGCTCTCCGCCGGGGCGCGGGCTGCCTCAACTT  
 CGCCGACTCCGCCGAGCTGCTC ACCGTGCCGGCGGCTCTCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTCTG  
 TGGAGGCCGTCGAGGACTTCTGCGGCGGGAGGCGATCGCCGAGGAGGACGCGCTCTCGGGCACCTCGTCTCCGCG  
 CCTCTCTCCCTCACCGACGACGAGTCTCTCTTCCCCGCGGAGGACTCGCCGTTTCGAGCTGGATGTCTGAGCGA  
 CATGGGCTGGGACCTGTACTACGCGAGCCTGGCGCAGGCGATGCTGATGGCGCGCCCTTCTTCCATGGCCGCGGCGC  
 TCGGCGATTACGGCGAGGTCGATGTGCCACTCTGGAGCTACCAGAGCTAGCTAGTTTCGCGCCAGTTCGAATAT  
 TTTAC

>TaCBFIIIIa-6.1\_N33

GCCAAGGAACAGAGCATCGAAGGAA GAAGAAG ATG TGTCCGATCAAGAAGGAGATGAGCGGGGAGTCCGGCTCGCCG  
 TGCAGCGCGGAGAACTTCTACTCGCCCTCCACGTCGCGGGGAGCATCAGCAGGCGAAGCAGGCGGCGTGGACGTCCGC  
 GCCGCGAAGCGGCCGGCGGGGCGGACCAAGTTTCAGGGAGACGCGGCACCCGGTATACCGCGGCGTGCAGCGCAGGG  
 GCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCGAC  
 ACCGCCGAGGCCGCCGCGCGCCAACGACGCGCCCATGATCGCGCTCTCCGCCGGGGCGCGGGCTGCCTCAACTT  
 CGCCGACTCCGCCGAGCTGCTCGCCGTGCCGGCGGCTCTCTCGTACCGCAGCCTCGACGAGGTCCGCCACGCCGTCTG  
 TGGAGGCCGTCGAGGACTTCTGCGGCGGGAGGCGATCGCCGAGGAGGACGCGCTCTCGGGCACCTCGTCTCCGCG  
 CCTCTCTCCCTCACCGACGACGAGTCTCTCTTCCCCGCGGAGGACTCGCCGTTTCGAGCTGGATGTCTGAGCGA  
 CATGGGCTGGGACCTGTACTACGCGAGCCTGGCGCAGGCGATGCTGATGGCGCGCCCTTCTTCCATGGCCGCGGCGC  
 TCGGCGATTACGGCGAGGTCGATGTGCCACTCTGGAGCTACCAGAGCTAGCTAGTTTCGCGCCAGTTCGAATAT  
 TTTAC

>TaCBFIIIIa-6.2\_N38

ACAAACACAGCTTCCGTTGGACAGCAAG ATG TGTCCGATCAAGAGGGAGATGAGCGGGGAGTCCGGCTCGCCGTCG  
 CGTGACGCGGGAGAACTTCTGCTCGCCCTCGCGCTCGCCGAGCGCCAGCAGGCGAGGCGGGGTGGACGTCTG  
 GCGCCGGCGAAGCGGCCGGCGGGGCGGACCAAGTTTCAGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCAGCGCAG  
 GGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGCAGGCGCGGGAGCAGGCTCTGGCTCGGCACCTTCG  
 ACACCGCCGAGGCCGCCGCGCGCCAACGACGCGCTCATGCTCATGCTCGCCGCCGAGGCGCCGCTGCCTCAAC  
 TTCCGCGACTCGGCCGAGCTGCTCTCCGTGCCGGTGGCTCTCTTACCAGCAGCCTGGACGAGGTCCGCCACGCCGT  
 CGTGGAGGCCGTGGAGGACTTGTGCGGCGGAGGCGCTCGCCGAGGAGGACGCGCTCTCGGGCACCTCTCTGTCG  
 CGCCCTCCCCCTCACCGACGACGAGTCTCTCATCGCCGTGCCGAGGAGGACTCGCCGTTTCGAGCAGGACGTG  
 CTGAGCGAGATGGGCTGGGACCTGTACTACGCGAGCCTGGCGCAGGCGATGCTCATGGCGCGCCCGCGCGGCTG  
 GCGCTCGGCGATTACGGGAGGCTACCTCGCCGACGTCGCACTCTGGAGCTACCAGAGCTAGCGCG TAGTTCGC  
 GCTGCTTCGAGTTT TAGCT

>TaCBFIIIIb-23.1\_S197 (manque\_154b\_pour\_ATG)

GCAGCGCCCCCGCTGGGCTCTGCGAGTGGCGCTGCGGCGAGCGCGGCACGCGCCTCTGGCTCGGGACGTACATC  
 ACGGCCGAGGCGGCCGCGGTGCGCACGACGCCCATGCTCATGCTCCGGGGCCGCTCCGCCGCGTGCTCAACTT



TCTCACTCCAGCGCAATCATCAAGCTAGACACTAGCTATCGACCGGCAGCCAGCAGCTATGGACATGAGCCTCGAGC  
ACTCGAGCTCCCCCTCCTCCTCATCCACTACCGAGCGCGCCGGGACGGCGTGGCCGTGGCCCGCCGAAGCGCCCCGCG  
GGCCGCACCAAGTTCCGGGAGACGCGGCACCCCGTGTTCCCGCGCGTGGCCCGCCGTGGCAGCGCCGGCCGGTGGGT  
CTGCGAGGTGCGCGTGCCCGGGGAGCGCGGCACGCGCCTCTGGCTCGGGACGTACATCACGGCCGAGGCGGCCGCGC  
GAGCGCACGACGCCGCCATGCTCATGCTCCGGGGCGCTCCGCCGCGTGCTCAACTTCCGGGACTCCGCGTGCGCTG  
CTCGCGGTGCCGCCCGCATTACCAACCTCTCTGACGTCCGGCGCGCGGCCGTCCAGGCCGTGCGGGACTTCTTGCG  
CCGTCCGGAGGGCCAAAGTGCCGCTGCCGCGGTGCAGGAGGTACCTCCAGCGTGTCCGCCCGCTCGTCGGCGGTGT  
GCAGTGTGCCCTCGTCAGAGACGGCGCAGGCTTCCCTGATGCCGCTTTTGAGGCGCCGGCCGCACTGGACATGGAC  
GTGTTTGACCTCGACTGCTTGTGTTGGGAAACGGACTCCGACGCGTACTACTACGGAACCTTGCGCAGGGGCTGCT  
CATGGAGCCACCGCGACTATGGCCACCGGGTCTACTGGGACAATGGAGACTGCGCCGACGGCGGAGCCGGAGCTG  
ATGTCGCGCTCTGG

TCTCACTCCAGCGCAATCATCAAGCTAGACACTAGCTATCGACCGGCAGCCAGCAGCTATGGACATGAGCCTCGAGC  
ACTCGAGCTCTCCCTCCTCCTCATCCACTACCGAGCGCGCCGGACGGCGTGGCCGTGCGCGCGAAGCGCCCCGCGG  
GCCGCACCAAGTTCCGGGAGACGCGGCACCCCGTGTTCCGCGCGGTGCGCCGCCGTGGCAGCGCCGGCCGGTGGGTG  
TGCAGGTTGCGCGTGCCCGGGGAGCGCGGCACGCGCCTCTGGCTCGGGACGTACATCACGGCCGAGGCGCCGCGCG  
AGCGCACGACGCCGCCATGCTCATGCTCCGGGGCCGCTCCGCCGCGTGCCTCAACTTCCGGGACTCCGCGTGGCTGC  
TCGCGGTGCCGCCCGCATTACCAACCTCTCTGACGTCCGGCGCGCGGCCGTCCAGGCCGTGCGGGACTTCTTGC  
CGTCCGGAGGCCAACAGTGCCGCTGCCGCGGTGCAGGAGGTCACTTCCAGCGTGTCCGCCCCGTCTGTCGGCGGTGTG  
CAGTGTGCCCTCGTCCAGAGCGGCGAGGCTTCCCCTGATGCCGCTTTTGAGGCGCCGGCCGCACTGGACATGGACG  
TGTTTGACCTTCGACTGCTTGTTTGGGGAAACGGACTCCGACGCGTACTACTACGCGAACCTTGCGCAGGGGCTGCT  
ATGGAGCCACCGCCGACTGTGGCCACCGGTCGTACT

TCTCACTCCAGCGCAATCATCAAGCTAGACACTAGCTATCGA.CCAGCAGCCAGCAGCTATGGACATGAGCCTCGAGC  
 ACTCGAGCTCTCCCTCCTCCTCATCCACGACCGAGCGCGCGGAACGGCGTGGCCGT~~~~~CCGAAGCGCCCGCT  
 GGCCGCACTAAGTTCCGGGAGACGCGGCACCCCGTGTTCCGCGCGCTGCGCCCGCTGGCAGCGCCGCCGGTGGGT  
 CTGCGAGGTGCGCGTGCCCGGGGAGCGCGGCACGCGCCTCTGGCTCGGGACGTACATCACGGCCGAGGCGGCCGCGC  
 GCGCACACGACGCGGCCATGCTCATGCTCCGGGGCCGCTCCGCCGCGTGCTCAACTTCCGGGACTCCGCGTGGCTA  
 CTCGCGGTGCCGCCCGCGTTACCAACCTCTCTGACGTCCGGCGCGCGGCCGCTCCAGGCCGTGCGGGACTTCTGCG  
 CCGTCCGGAGGCCACCAAGTGCCCGTGCCCGGGTGCAAGGAGTACCTCTCAGCGTGTCGCCCCCGCTCGTCGGCAGTGT  
 GTAGTGTGCCCTCGTCAGAGACGGCGAGGCTTCCGTGAGTACGCCCTTTTAGGCGCCGCCGCGACTGGACATGGAC  
 GTGTTGCACCTTGACTGCTTGTTTGGGGAACGGATTCCGATCGCTACTACGCGAACCTTGCGCAGGGGCTGCT  
 CATGGAGCCACCGCGACTCTGGCCACCGGGTCTACTGGGACAATGG

TCCGGACGTTACATACCGGCGAGGCGCGCGCGCACAGACGCCGCATGCTCATGCTCCGGGGCCGCTCCGCC  
GCGTGCTCAACTTCCGGGACTCCGCGTGGCTGCTCGCGGTGCCGCCGCGTTACCAACCTCTCTGAGCTCCGGCG  
CGCGGCCGCTCAGAGCCGTCGCGGACTTCTGCGCCGTCGCGAGGCCACAGTGCCTGCGCGGTGCGAGGATCA  
CCTCCAGCGTGTCGCGCCCGCTCGTGGCAGTGTGTAGTGTGCCCTCGTCAGAGACGGCGAGGCTTCCCTGATGCC  
GCTTTTGAGGCGCCGGCCGCACCTGGACATGGACGTGTTGCACCTTGACTGCTTGTTTGGGGAAACGGAATCCGATGC

GTACTACTACGCGAACCTTGCGCAGGGGCTGCTCATGGAGCCACCGCCGACTCTGGCCACCGGGTCTGACTG  
ACAATGGAGACTGCGCCGACGGCGGAGCCGAGCTGATGTCGCGCTCTGGAGTTACTAGTATTACTAATATCAACA  
TCTCCAACATATATTTCGAGTTGAAGCCGCTTCGATTATGATCCACGTCCCGAGAGCTG

>TaCBFIIb-23.2.1-N5

ACTAGCTATCGACCAGCAGCCAGCAGTATGGACATGAGCCTCGAGCACTCGAGCTCTCCCTCCTCCTCATCCACGA  
CCGAGCGCGCCGGAACGGCGTGGCCGTGGCCGCCGAAGCGCCCCGCTGGCCGCACTAAGTTCCGGGAGACGCGGCAC  
CCCGTGTTCCGCGGCGTGCGCCGCCGTGGCAGCGCCGCGCGGTGGGTCTGCGAGGTGCGCGTGCCCGGGAGCGCGG  
CACGCGCCTCTGGCTCGGGACGTACATCACGGCCGAGGCGGCGCGCGCACACGACGCCGCCATGCTCATGCTCC  
GGGGCCGCTCCGCCGCGTGCTCAACTTCCGGGACTCCGCGTGGCTGCTCGCGGTGCCGCCCGGCTTACCAACCTC  
TCTGACGTCCGGCGCGCGGCCGTCCAGGCCGTGCGGCACTTCTGCGCGTCCGGAGGCCACAGTGCCGCTGCCGC  
GGTGACGAGGTACCTCCAGCGTGTCCGCCCGCTGCTCGGCAGTGTGTAGTGTGCCCTCGTCAGAGACGGCGCAGG  
CTTCCGCTGATGCGCGTCTTTGAGGCGCCGCCACACTGGACATGGACGTGTTGACCTTGACTGCTTGTGTTGGGAA  
ACGGATTCCGATGCGTACTACTACGCGAACCTTGCGCAGGGGCTGCTCATGGAGCCACCGCCGACTCTGGCCACCGE  
GTCGTACTGGGACAATGGAGACTGCGCCGACGGCGGAGCCGAGCTGATGTCGCGCTCTGGAGTTACTAGTATTA

>TaCBFIIb-23.3-M24 (manque 208b pour\_ATG)

TCTGGCTCGGGACGTACATCACGGCCGAGGCGGCCGCGCGCGCACGACGCCGCCATGCTCATGCTCCGCGGCCGC  
TCCGCCGCGTGCTCAACTTCCGGGACTCCGCGTGGCTGCTCGCGGTGCCGCCGCGTTACCAACCTCTCTGACGT  
CCGGCGCGCGGCCGTCCAGGCCGTGCGGCACTTACTGCGCGTCCGGAGGCCACAGTGCCGCTGCCGCGGTGCAAG  
AGGTACCTCCAGCGTGTCCGCCCGCTGCTCGGCGGTGTGCAGTGTGCCCTTGTGAGAGACGGCGCAGGCTTCCGC  
GATGCCGCTTTTGGGCGCGGCTGCACTGGACATGGACGTGTTGACCTCGACTGCTGTTTGGGGAAACGCACT  
CGACACGTACTACTACGCCAACCTTGCGCAGGGGCTGCTCATGGAGCCACCGTCGACTGTGGCCACCGGGTCTGTA  
GGGACAATGGAGACCGCGCCGACGGCGGAGCCGGAGCTGATGTCGCGCTCTGGAGTTACTAGTACTAATATCAACA  
TCTCCAACATATATTTCGAGTTGAAGCCGCTTCGATTATGATCCACATCTCGAGACAGTCCATCGATGGCGCGCGCA  
TCAAATTGGATGGAATGGATGATGTTCTGCAATTTCTTCTCAGGTTCTGCGTGTCTCCTCATTCTCTCAAAGGG  
CGAATTCCAGCACACTGCGGCCGTTACTAGTGATCCGAGCTCGGTACCAAGCTTGATGCATAGCTTGAGTATTCT  
ATAGTGTACCTAAATAGCTGGCGTAATCATGGTCA

>TaCBFIIc-3.1-M159 (manque 132b pour\_TAG)

ATGATTACGCCAGCTATTTAGCTCACACTATACAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGATCCA  
CTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTACAATTAGCTCACGGCCTCACACTCCTCAGCCTCAGTAAGC  
TCAAGCACCAAGCTCGACTGCTCAAGCAGGAATCCACCAGTCAATCACCGAGCACTCAGCCGGTAGCCATGGACATG  
GGCCTTGAGGTCTCGAGCTCCTCCCGTCTCTCTGTCGCTGCTGCTCTCGCCCGAGCACGCGGCGGCGCGCGCTC  
GCTGGCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGACACCCGGTGTACCGCGGCGTGCGGCGCGGG  
GCAACGCCGAACGTGGGTCTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGGAGGCTCTGGCTCGGGACGTACGAC  
ACGGCTGAGCTCGCAGCGCGCGGAACGACGCCGCCATGCTTGCCCTGGGCGGCCGCTCCGCCGCGTGCCCTCAACT  
CGCGGACTCCGCGTGGTTGCTCGCCGTGCGGTCCGCACTCTCCGACCTCGGCGACGTCCGGCGCGCGGCGGTGAGG  
CCGTGGCGAACTTGACGCGACGAAAGGCTGGCAACGGCTCCCTCACCGCCACCGTCACCGAAGAGGCCCTCTGTGAC  
GCTCTGAAGAATCGTCGTCTGAGTCTGACAGTGCCGGTTCGTCGGAGACGTCGGAACCTTCAGCCGATAGAGAGTT  
CGAGGTGCCGCTCGAGTGACACCGATATGTTCCGGCTTGACTTGTTC

>TaCBFIIc-D3\_S153 (INDEL)

ACAATTAGCTCACGGCCTCACACTCCTCACTGAGCTCAACCACCAAGCTCGACTGCTCAAGAAGAAGGCCACCTGCC  
AATCACCCAGCACTCTGCCGCTAGCCATGGACATGGGCCTTGAGGTCTCGAGCTCCTCCCCATCTCTCTGTCGGTG  
TCGTCTCGCCGTCACGCGGCGGCGCGCGCTGCTGGCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGAC  
GCGGCACCCGGTGTACCGCGGAGTGCGGCGCCGAGGCAACGCCGAACGGTGGGTCTGCGAGGTGCGCGTCCCCGGCA  
AGCGCGGCGCGAGGCTCTGGCTCGGGACGTATGCCACGGCTGAGGTGCGCGCGCGCGGAACGACGCCGCCATGCTT  
GCCCTGGGCGGCGCTCCGCCGCTGCTCAACTTCCGGGACTCCGCGTGGCTGCTCGCCGTGCCGCCCGCACTCTC  
CGACCTCGGCGAGTCCGGCGCGCGGCGGTGAGGCCGTGCGGATTTCCAGAGGCGAGAGGCTGCCAACGGCTCC  
TCACCGCCACCGTCACCGAAGAGGCCCTCTGTGGCGCTCTGAAGAATCGTCGTCTGAGTCTGACAGTGCCGGTTCG  
TCGGAGACGTCGGAACCTTCTGCCGATGACAGAGTTCGAGGTGCTGTCGAGTGACACCGATATGTTAGTAGGCT



TGACTTGTTCCTCCGAAATGGACCTGTGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCGTGGACCCGCCGT  
CGACGGTGGCCATCATCGACTCGTACTGGGACAACGGCGACGACGAGCTGATGTGCGCTCTGGAGCTAC TAG CTA  
GTACTTAGTGTCGATAATCCCCCGCAAAA

>TaCBFIIIc-D3\_S154 (manque\_49b\_pour\_ATG)

TGTCGTCCTCGCCCGTGTACGCGGCGGGCGGGCGTCTGCTGGCCAAGCGCCCGCGGGGCGCACCAAGTTCGGGAG  
ACGCGGCACCCGGTGTACCGCGGAGTGCAGCGCCGAGGCAACGCCAAGCGGTGGGTCTGCGAGGTGCGCGTCCCCGG  
CAAGCGCGGCGCGAGGCTCTGGCTCGGGACGTACGCCACGGCTGAGGTGCGCCGCGCGCGCAACGACGCCGCCATGC  
TTGCCCTGGGCGGCGCTCCGCCACGTGCTCAACTTCGCGGACTCCGCGTGGCTGCTCGCCGTGCCGCCGCACTC  
TCCGACCTCGGCGACGTCCGGCGCGCGCGGCTCGAGGCCGTTCGCGGATTTCCAGAGGCGAGAGGCTGCCAACGGCTC  
CCTCACCGCCACCGTCACCAAGAGGCCCTCTGTGGCGCTCTGAAGAATCGTCGCTGAGTCTGACAGTGCCGGTT  
CGTCGGAGACGTGGAACTTCTGCGGATGACAGAGTTCGAGGTGCTGTCGAGTGGACACCGATATGTTTCACTAGG  
CTTGACTTGTTCCTCCGAAATGGACCTGTGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCGTGGACCCGCCGT  
GACGGTGGCCATCATCGACTCGTACTGGGACAACGGCGACGACGAGCTGATGTGCGCTCTGGAGCTAC TAG CTAG  
TACTAGTGTGATAATCCCCCGCAAAA

>TaCBFIIIc-10.1\_S160 (manque\_44b\_pour\_ATG)

CAACGAGAACCGCTTGGTCCCAAGCGCCCGCGGGGCGCACCAAGTTCGGGAGACAAGGCACCCGGTGTACCGCG  
GCGTGCAGCGCGCGGGCAACGCCGAGCGGTGGGTCTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGCGAGGCTCTGG  
CTCGGGACGTATGCCACGGCCGAGATCGCAGCGCGCGCAACGACGCCGCCATGCTCGCCCTGGGCGGCGCTCCGC  
CGCGCGCCTTAACCTCCCGGACTCCGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGC  
ACGCGGCGGTTCGAGGCCGTTCGCGGATTGACGCGACGGGAGGCTGCCGGTGGGTCCATCACCGCCACCGTCAACGAG  
GAGGCCTCTGTGGCGCTCTGCGGAGTCTGCTCGGAGTCTGACGATGCCGTTTCGTGCGAGACGTGAAACCTTC  
TGCGCATGGGCGCTTTGCGCTGCCGGGCGGAATGGACGTGAAATGTTTCACTAGGCTTACTTGTTCCTCCGAAACGG  
ACTTGGGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCATGGACCCACCGCCGTTGGCGACCGGCACCGGCGCG  
TACTGGGATGACGGAGAGTTCGGCGAGGTGGCAACTGAGTTCGCGCTCTGGAGCTTG TAG CCGATTCTGCTATGTTT  
TGACTCTGTAGTCTTTATTTTTCAGTTGAGAATTCATTATAACACAGTTTTATTCCACTGTTTATACTCCGGTTC  
TTCTCTAGCTGAGACG

>TaCBFIIIc-10.1-M42 (manque\_63b\_pour\_ATG)

GCCACGCCCCCGCGCGGGCGCACCAAGTTCGGGAGACAAGGCACCCGGTGTACCGCGGCGTGCAGGCGACCGGGCAA  
CGCCGAGCGGTGGGTCTGCGAGGTGCGCGTCCCCGGCAAGCGCGCGGAGGCTCTGGCTCGGGACGTATGCCACGG  
CTGAGATCGCAGCGCGCGCAACGACGCCCGCATGCTCGCCCTGGGCGGCGCTCCGCCGCGCGCCTTAACCTCCCG  
GACTCCGCGTGGCTGCTCGCCGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGCGCGCGGCGGTTCGAGGCCGT  
CGCGGATTTGACGCGACGGGAGGCTGCCGGTGGGTCCATCACCGCCACCGTCAACGAGGAGGCCTCTGTGGCGCTC  
CTGCGGAGTCTGCTCGGAGTCTGACGATGCCGTTTCGTGCGAGACGTGAAACCTTCTGCGCATGGCGCTTTGCG  
CTGCCGGGCGGAATGGACGTGAAATGTTTCACTAGGCTTACTTGTTCCTCCGAAACGGAATTGGGCTCTGCTACGC  
GAGCCTCGCGGAGGCGCTGCTCATGGACCCACCGCGGTGGCGACCGGCACCGGCGCGTACTGGGATGACGGAGAGT  
TCGGCGAGGTGGCAACTGAGTTCGCGCTCTGGAGCTTG TAG CCGATTCTGCTATGTTTGTACTCTGTAGTCTTTATT  
TTTTCAGTTGAGAATTCATTATAACACAGTTTTATTCCACTGTTTATACTCCGGTCTTCTCTAGCTGAGACG

>TaCBFIIIc-B10\_N73

CAAGGAAGCCACCTGCCAATCAC CAGCACTCTGCCGGTAGCATGACATGGGCCTTGAGGTCTCGAGCTCCTCCC  
CATCTCTCTCTCTCTGCGCCACCGCCCGCGGGGCGCACCAAGTTCGCGAGACCGGCACCCCTGTGTATCGTGGC  
GTGCGGCGCGGGGCAACGCCAACGGTGGGTCTGCGAGGTGCGCGTCCCTGGCAAGCGCGGCGCTCGGCTCTGGCT  
CGGGACTTACGCCACGGCCGAGATTGACGCGCGCGCAACGACGCCGCCATGCTCGCCCTGGGCGGCGCTCCGCCG  
CGCTCCTCAACTTCCCGGACTCCGCGTGGCTGCTCGCTGTGCCGTCCGCGCACTCCGATCTCGCCGACGTCCGGCGC  
GCGGCGGTTCGAGGCCGTTCGCGGATTTGACGCGACGGGAGGCTGCCGGTGGGTCCATCACCGCCACCGCATCGCCAC  
CGCCCGCGAGGAGGCTCTGTGGCGCTCTGACAGATCGTCTGAGTCTGACGATGCCGGTTCGTGCGAGACGT  
CGAAACCTTCCGCCGATGGAGACTTCGCGGTGCCGGGCGGAATGGACATCGAAATGTTTCACTAGGCTTACTTGTTC  
CCGGAATGGACTTGGGCTCGTACTACGCGAGCCTCGCGGAGGCGCTGCTCATGGACCCGCTCCGGTGGCGACTGG





CCATAGATCGACCTTGGCTACCGTACGTAAGCTCGAGCCAGCACCGACACCTGCAGTAGCAATGGACATGAC  
CGCCTCCCATCAGCAGCGGAGCTCCCCGTCTCGCCGTCTCTCTCTCGCACCTGAAGCGTCCCCCGGGCGCACCA  
AGTTCAAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGGCGCCGCGGACGCGCCGGCCGGTGGGTGTGCGAGGTG  
CGCGTCCCCGGCAAGCGCGGCGAGCGGCTGTGGCTCGGCACGCACCTCACTGCCGAGGCGGCCACGCGAGCGTACGA  
CGCCGCCATGCTCTGCCTGATCGGCCCCCTCCACCCAGTGTCTCAACTTCGCCGACTCCGCCTGGCTCCTCGCGGTGC  
CGTCCGCGCTCCCCGACTTTGCAGACGTCCGGCGCGCGGCTCTCAGCGCTGTGCGAGACTTCCAGCGCCGGGAGGCT  
GCCAGCGGCGCCGCCACAGGTCCCTGGACGCCACCGTTCCCGTCGATGATGGAACCTGTAGCCAATCCGCGCAGTC  
GTCCATGGAAAATACCGGCTCATCATGGACGTCGTCTCCCTGCCGTCTGGCAACGGAATGTTTGAGGTGCCGCCA  
CACTAGGCTGCGACATGTTTCGAGCTGGATATGTCCGGGAAATGGACCTGGACACATACTATGCGTACTTCGCGGAG  
GGGCTTCTCTCGAGCCGCCGAACCGCCGGTCCGCGGCGGTGCTGG

>TaCBFIIId-24.2b\_S68 (manque\_23b\_pour\_ATG)

GATCACTTCCCTTCTCATCGACGTCCGGGCACGAGCACGGGAGGTGGTGCTTGTGTGGTCGCCGGCGGCGAAGC  
GGCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCAGCGCCGGGGCAGTCCGGGG  
CGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGCTCGGGACGTACGTCCCGCCGAGTC  
GGCCACGCGCGCGCACGACGCCACCATGCTCGCGTGTCTCGGCCACTCCGCCTCCGCCCGGGCGTGCCTCAACTTCC  
CGGACTCCGCGTGGTGTCTGTCATGCCCCGTGGCTCTCCGACTTGGCCGATATCCGACGCGCGGCCATCGAGGCC  
GTCCGATTTTCTGTGCTGGAGGCCGCCCGCGTGGTGCCCATCATCGACGAGGCCACCTCGCCCGTGTACCTGCC  
GTCGCGCGTGGACAATGCCGACGAAGTGTTCAGGTGCCGACTTCTCTGCGCAGAGCAGCGACATGTTCAAGCTCG  
ACATGTCCGGGAAATGGACCTGGACGCGTACTACGACGGCTTTGCGCAGGGGATGCTCCTCGAGCCGCCGCCACG  
CCGACGTACTGGGAAAACGGAGAATGCGGCGACGGCGGAGCGGCCGCCGCCCTCTGGAGCTACTGATGTGATAACC  
TCCAACGGGTTTCACTG

>TaCBFIIId 24.2b\_M78

GCAGCAATGGACATGGCCATCGACAGCTGGATCAGCTCCCTTCTCATCGACGTCCGGGCACGAGCACGGGGAGGT  
GGTGCTTGTGTGCTACCGCGGGCGAAGCGGCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACC  
GCGGCGTGCAGCGCCGGGGCAGTGCAGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTC  
TGGCTCGGGACGTACGTGCGCGCCGAGTCGCGCCACGCGCGCGCACGACGCCACCATGCTCGCGTGTCTCGGCCACTC  
CGCTCCGCCCGCGGCTGCCTCAACTTCCCGGACTCCGCGTGGCTGCTCGTCATGCCCCGTGGCTCTCCGACTTGG  
CCGATATCCGACGCGCGGCCATCGAGGCCGTGCGGATTTCTGTGCTGGAGGCCGCCCGCGTGGTGCCCATCATC  
GACGAGGCCACCTCGCCCGTGTACCTGCCGTGCCCCGTGGACAATGCCGACGAAGTGTTCAGGTGCCGACTTCTC  
TGCGCAGAGCAGCGACATGTTTCGAGCTCGACATGTCCGGGAAATGGACCTGGACGCGTACTACGACGCTTTGCGC  
AGGGGATGCTCTCTCGAGCCGCCGCCACGCGACGTAAGTGGGAAAACGGAGAATGCGGCGACGGCGGAGCGGCCGCC  
GGCTCTGGAGCTACTGATGTGATAACCTCCAACGGGTTTCACTG

>TaCBFIIId 24.2b\_M80 (manque\_4b\_pour\_ATG)

ACATGGCCATCGACAGCTCGATCAGCTCCCTTCTCATCGACGTCCGGGCACGAGCACGGGGAGGTGGTGCTTGTG  
TGGTCGCCGGCGGCGAAGCGGCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGCGCGTGC  
GCGCCGGGGCAGTGCAGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGCAAGCGCGGCGAGCGGCTCTGGCTCGGGA  
CGTACGTGCGCGCCGAGTCGGCCACGCGCGCGCACGACGCCACCATGCTCGCGTGTCTCGGCCACTCCGCCTCCGCC  
GCGGCGTGCCTCAACTTCCCGGACTCCGCGTGGCTGCTCGTCATGCCCCGTGGCTCTCCGACTTGGCCGATATCCG  
ACGCGCGGCCATCGAGGCCGTGCGGATTTCTGTGCTGGAGGCCGCCCGCGTGGTGCCCATCATCGACGAGGCCA  
CCTCGCCCGTGTACCTGCCGTGCGCCGTGGACAATGCCGACGAAGTGTTCAGGTGCCGACTTCTCTGAGCAGAGC  
AGCGACATGTTTCGAGCTCGACATGTCCGGGAAATGGACCTGGACGCGTACTACGACGGCTTTGCGCAGGGGATGCT  
CCTCGAGCCGCCGCCACGCGACGTAAGTGGGAAAACGGAGAATGCGGCGACGGCGGAGCGGCCGCCGCCCTCTGGA  
GCTACTGATGTGATAACCTCCAACGGGTTTCACTG

>TaCBFIVa-A2\_S70

TCGTTACACCACAAACCACTCTCAACGCCAGCTGCGACCGATGGACACCAACGCCGCTGGCCGAGTTTGACGGGG  
AAGAGTACAGGACGGTGTGGCCGAGGACGAGTACCGGACGGTGTGGTTCGAGCCGCCGAAGCGGCGGGCGGGG  
CGGAACAAGTTGACAGGACACGCCACCCAGTGTACCGCGCGTGCGCCCGGTGGCCGGGAAGGGCAGTGGGTGTG  
CGAGCTGCGCGTCCGGCCGGAAGCCGGATTACTCCAGGATCTGGCTTGGACCTTCGCCAGTGCCAGATGGCGG



CGCGCGCGCAGACTCGGCCGCGCTCGCGCTCTCCGGCCGCGACGCGTGCCTCAACTTCGCCGACTCCGCCTG  
 CGGATGATGCCCGTCCACGCAGCCGGGTCTTCAAGTTGGCCGCCGCGCAGGAGATCAAGGACGCCGTGCGCGTGG  
 CCTCAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGACGAGTCAACGGCGCCGTCTGCCACGGCCGAGGAGAGCGCG  
 TCTCCATCATCCCCAGCGACCTGTGCGGGCTCGACAATGAGCACTGGATCGGCGGCATGGAGGCCGGGTGCGCACT  
 GCGAGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGGCCGACGGAGCTTGGCGGGAGGACCGCGAACACGACGACCG  
 ATTTCGACACGTCGCTGTGGAGCTACTAGTGTGATCAACTGATTAAGCAATGTAAAGATCTAGAGAGTACTGCTAGTG

>TaCBFIVa-A2\_S71

TCGTTACACCACAAACCACTCTCAACGCCAGCTGCGACCGATGGACACCAACGCCGCCTGGCCGCGAGTTTGACGGG  
 AAGAGTACAGGACCGGTGTGCGCCGAGGACAGGAGTACCGGACCGGTGTGGTTCGGAGCCGCCGAAGCGGCGGGCGGGG  
 CGGAACAAGTTGCAGGAGACACGCCACCCAGTGTACCGCGGCTGCGCCGCCGTGGCCGGGAAGGGCAGTGGGTGT  
 CGAGCTGCGCGTGC CGGCCGGAAGCCGAGTTACTCCAGGATCTGGCTTGGCACCTTCGCCAGTGCCAGATGGCGG  
 CGCGCGCGCAGACTCGGCCGCGCTCGCGCTCTCCGGCCGCGACGCGTGCCTCAACTTCGCCGACTCCGCCTGGCGG  
 ATGATGCCCGTCCACGCAGCCGGGTCTTCAAGTTGGCCGCCGCGCAGGAGATCAAGGACGCCGTGCGCGTGGCCCT  
 CAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGACGAGTCAACGGCGCCGTCTGCCACGGCCGAGGAGAGCGCGCTCT  
 CCATCATCCCCAGCGACCTGTGCGGGCTCGACAATGAGCACTGGATCGGCGGCATGGAGGCCGGGTCTACTACGCG  
 AGCTTGGCGCAGGGGATGCTCATGGAGCCGCCGGCCGACGGAGCTTGGCGGGTGACCGCGAACACGACGACGGAT  
 CAACACGTCGCTGTGGAGCTACTAGTGTGATCAACTGATTAAGCAATGTAAAGATCTAGAGAGTACTGCTAGTGCA  
 GGGGCGAATTCTGCCTGC

>TaCBFIVa-A2-M8

CAACGCCAAGTCCGACCGATGGACACCAACGCCGCCTGGCCGCGAGTTTGACGGGCAAGAGTACAGGACCGGTGTGGCC  
 GGAGGAGCAGGAGTACCGGACCGGTGTGGTTCGGAGCCGCCGAAGCGGCGGGCGGGGCGGAACAAGTTGCAGGAGACAC  
 GCCACCCAGTGTACCGCGGCGTGC CGGCCCGGTGGCCGGGAAGGGCAGTGGGTGTGCGAGCTGCGCGTGC CGGCCCGG  
 AGCCGGAGTTACTCCAGGATCTGGCTTGGCACCTTCGCCAGTGCCAGATGGCGGCGCGCGCAGACTCGGCCGC  
 GCTCGCGCTCTCCGGCCGCGACGCGTGCCTCAACTTCGCCGACTCCGCCTGGCGGATGATGCCCGTCCACGCAGCCG  
 GGTGCTTCAAGTTGGCCGCCGCGCAGGAGATCAAGGACGCCGTGCGCGTGGCCCTCAAGGAGTTCCAGGAGCAGCAG  
 CGCCCTGCCGACGAGTCAACGGCGCCGTCTGTCACGGCCGAGGAGAGCGCGCTCTCCATCATCCCCAGCGACCTGT  
 GGGGCTCGACAATGAGCACTGGATCGGCGGCATGGAGGCCGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCA  
 TGGAGCCGCCGGCCGACGGAGCTTGGCGGGAGGACCGCGAACACGACGACGCGATTTCGACACGTCGCTGTGGAGCTAC  
 TAGTGTGATCAACTGATTAAGCAATGTAAAGATCTAGAGAGTACTGCTAGTGCAAGGGCGAATTCCAGCACACTGGC  
 GGCGGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGATGCATAGCTTGAGTATTCTATAGTGTACCTAAATAGC

>TaCBFIVa-25.1\_S97 (manque\_32b\_pour\_ATG)

GCCGGAACAACTTCCAGCGACCGCCCGCCCGGCTTACCGCGCGGCTTCGCCGCCGTGGCCGGGAAGGGCAGTGGGTG  
 TGGGTGTGCGAGCTGCGGGTCTGGCCGCCGGGAGCCGGGTTTACTCCAGGATCTGGCTCGGCACCTTCGCCGACCC  
 CGAGATGGCGGCGCGCGCGCTGACTCGGCCGCGCTCGCGCTCTCCGGCCGCGATGCGTGCTCAACTTCGCCGACT  
 CCGCCTGGCGGATGATGCCCGTCCACGCGGCCGGGTCTTCAAGCTGGCCGCCGCGCAGGAGATCAAGGACGCCGTG  
 GCCGTGCGCCTCAAGGAGTTCCAGGAGCAGCAGCGCCCTGCCGACGTGTCAAAGGCGCCATCTCCACGGACTCCAC  
 GTCCGAGGAGAGCGCGCCGTCCATCACCTCCAGCGACTTGTCTGGTCTGGACGACGAGCACTGGATCGGCGGCATGG  
 ACGCCGGGTCTACTACGCAAACTTGGCGCAGGGCATGCTCATGAGCCGCCGCCGCGCGGAGCGTGGCGGGAGGAC  
 CGCGAACAGGACGACGCGTGCACACGTCGCTGTGGAGCTACTAGTGTGACGATTGGATGTGTAAAGGTTGTAGA  
 GAGCAGTACTGTGTT

>TaCBFIId- A19a\_N97

GCTCAAGCTCAACTAGAGTGACCACATTCCAACCAAGCGATCGGCACTCCGGCAGCTAGCGATGACTTTGGTATCAA  
 CGGCTCGATCACTCCCTTCTCTCATCGAGTCCGGGCACGAGCTCGGGGATGCGGTGCGGCTGGTTCGCCGGCGG  
 CGAAGCGGCCCCGAGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGTGGCGTGC GGCGCCGGGGCAGC  
 GCGGGCCGGTGGGTGTGCGAGGTGCGGTCCCCGGAAGCGCGGCGAGCAGCTCTGGCTCGGACGTACGTGCGCGC  
 CGAGTCCGCCGCGCGCGTACGACGCCGCCATGCTCGCCCTGCTCGGACGCTCCCCCTCCGCCGCGCGGCCCTCA  
 ACTTCCCGGACTCCGCGTGGTGTGTCATGCCCCCGAGGCTCTCCGACCTGGCCGACGTCCGGCGCGCGGCCATC



CAGGCCGTCGCGGGCTTCTGCGCCTGGAGGCCGCCACCGTGGTCCCGGACGTCGACGAGGCCACCTCCCCTGT  
GTACCTGCCGTCGCCCCGTGGACAATGCGGACGAAGTGTTCCAGGTGCCGACTTTCTCCCCGCTGGGCAGCGACATGT  
TCGAGCTCGACATGTCCGGGGAAATGGACCTGGACGCGTACTACGCGGGCTTTGCCAGGGGATGCTCCTGGAGCCG  
CCGCCACGCCAGCGTACTGGGAGACCGGAGAGTGC GGCGACGGCGGAGCGGCCGCCGCTCTGGAGCTACTGATG  
TCTGTAACCTCCAGTGGGTTC

>TaCBFIVb-A20\_M101 (manque\_66b\_pour\_TAA) (INDEL)

ATCCCAACTCCACCGCTCAAACCAACCTGCAACTCTCAAGGCAGCGACTTTCCGCTAGTTTTTGACGCTGCAACCGA  
TGACACCGCCGCCCGGCTCCCCGCGTGAGGGGACAGGACGGTGTGCTCCGAGCCGCCAAGAGGCTGGCAGCC  
CGACCAAGTTCAAGGAGACGCGCCACCGCTGTACCGCGGCGTGCGGCGCGGGGCCGGCTCGGGCAGTGGGTGTG  
CGAGGTGCGCGTGCGCGGCGCGCAAGGTTACAGGCTCTGGCTCGGCACCTTCACCACTACCGAGATGGCGGCGCGG  
CGCAGCACTCCGCGCTCTCGCGCTCTCGACACCGCGCTTGCTCAACTTCGCGGACTCCGCGTGGCGGATGCTG  
CCGTCCTCGCGGCTGGCTCGTCCGCTTCAGCAGCGCGCGGAGATCAAGGACGCCGTCGCCGTCGTGGAGTTCCA  
GCGGCAGCGGCCGCTCTGTCCACGCCGAGACGCACGACGGCGAAAAGGACGTCCAAGGCTCACCAGCGCCGCGG  
AGCTGTCCACGTCCAGCGACTTGTGGACGAGCACTGGTTTGGCGGCATGAACGCCGCTCGTACTATGCGAGCTTG  
GCGCAGGGGATGCTCATGGAGCCGCCGCGCCAGAGCGCG

>TaCBFIVc-14.1\_S107

GCAGCAGCTAAACTCCCTGCTTAATTACCCACAGTCGGCAGGCTCCCGGCGACACTGCGATCGATCATGAGCGCC  
GCTGATGCCCGCTCCCCGTGTGATGGGCAAGGACGGTGTGGTCGGAACCGCCGAAGCGGCCGGCCGACCA  
GTTCAAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCGACGGGGCCCCGCCGCGGCTGGGTGTGCGAGGTGC  
GCGTGCTCGGGATGAGGGGCTCCAGGCTTTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGCGCACAGC  
GCCGCGGTTCTCGCGCTCTCCGGCCGCGCGCTTGCTCAACTTCGCCGACTCCGCTGGCGGATGCTCCCCGTCT  
CGCCGGCCCGTTTACGACGGCCAAGGAGATCAAGGATGCCGTCGCCGTCGCGCTCTGGCGTTCCAAAGGCAGCACC  
CGGTGCGCTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCTTGCCGGCACC  
AGCGCCCTGTCCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCGTGTACTCGGA  
GTTTATGGAGTCGCGGACACCAGACCATGGCGGGAAGACTTTGAGCTCGGTGCGTTCGAGACACCGCCATGGAGCT  
ACTTGTTCGACTAAGCAGTTAAAGTTTTTAGACGGTTTTCTGCTTCTGTT

>TaCBFIVc-14.1\_M116

GCAGAATTGCGCCTTGACGAGCTAAACTCCCTGCTTAATTACCCACAGTCGGCAGGCTCCCGGCGACACTGCGAT  
CGATCGATGAGCGCCGCTGATGCCGCTCCCCGTGTGATGGGCACAGGACGGTGTGGTCGGAACCGCCGAAGCGGCC  
GGCGGGCTGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCGACGGGGCCCCGCCGCGG  
GGGTGTGCGAGGTGCGCGTGCTCGGGATGAGGGGCTCCAGGCTTTGGCTCGGCACCTTCACCACCGCCGAGATGGCA  
GCGCGCGCGCACGACGCCCGCTTCTCGCGCTCTCCGGCCGCGCGCTTGCTCAACTTCGCCGACTCCGCTGGCG  
GATGCTCCCCGTCTCGCCGGCCCGTTTACGACGGCCAAGGAGATCAAGGATGCCGTCGCCGTCGCCGTCCTGGCGT  
TCCAAAGGCAGCACC CGGTGCGCTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGAT  
GGCTTGCCGGCACCAGCGCCCTGTCCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGG  
ATCGTGCTACTCGGAGTTCATGGAGTCGCGGACACCAAGACCATGGCGGGAAGACTTTGAGCTCGGTGGCGTCGAGA  
CACCGCCATGGAGTACTTGTTCGACTAAGCAGTTAAAGTTTTTAGACGGTTTTCTGCTTCTGTTCCA

>TaCBFIVc 14.1\_M203

GCAGCAGCTAAACTCCCTGCTTAATTACCCACAGTCGGCAGGCTCCCGGCGACACTGCGATCGATCATGAGCGCC  
GCTGATGCCGCTCCCCGTGTGATGGGCACAGGACGGTGTGGTCGGAACCTGCCGAAGCGGCCGGCCGCTGGACCA  
GTTCAAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCGACGGGGCCCCGCCGCGGCTGGGTGTGCGAGGTGC  
GCGTGCTCGGGATGAGGGGCTCCAGGCTTTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGCGCACGAC  
GCCGCGGTTCTCGCGCTCTCCGGCCGCGCGCTTGCTCAACTTCGCCGACTCCGCTGGCGGATGCTCCCCGTCT  
CGCCGGCCCGTTTACGACGGCCAAGGAGATCAAGGATGCCGTCGCCGTCGCCGTCCTGGCGTTCCAAAGGCAGCACC  
CGGTGCGCTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCTTGCCGGCACC  
AGCGCCCTGTCCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCGTGTACTCGGA  
GTTTATGGAGTCGCGGACACCAAGACCATGGCGGGAAGACTTTGAGCTCGGTGGCGTCGAGACACCGCCATGGAGCT  
ACTTGTTCGACTAAGCAGTTAAAGTTTTTAGACGGTTTTCTGCTTCTGTTCCA

>TaCBFIVc-14.1\_M204

GCAGCAGCTAAACTCCCTGCTTAATTACC CCATAGTCGGCAGGCTCCCGGCGACACTGCGATCGATCGATGACGCC  
GCTGATCCCCCTCCCGTGTGATGGGCACAGGACGGTGTGGTCGGAACCGCCGAAGCGGCCGGCCGCTCCGACCAA  
GTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGC GGCGACGGGGCCCCCGCGCGCGGTGGGTGTGCGAGGTGC  
GCGTGCTCGGGATGAGGGGCTCCAGGCTTTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGCGCACGAC  
GCCGCCGTTCTCGCGCTCTCCGGCCGCGCGCTTGCTCAACTTCGCCGACTCCGCCTGGCGGATGCTCCCCGTCT  
CGCCGGCCCGTTTACGACGCGCAAGGAGATCAAGGATGCCGTGCGCGTCCGCGTCTGGCGTTCCAAAGGCAGCACC  
CGTTCGCGTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCTTGCCGGCACCG  
AGCGCCCTGTCCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCGTGCTACTCGGA  
GTTTCATGGAGTCGCCGGATACAGACCATGGCGGGAAGACTTTGAGCTCGGTGGCGTCGAGACACCGCCATGGAGCT  
ACTTGTTTCGACTAAGCAGT

>TaCBFIVc-14.1\_N139

GCAGCAGCTAAACTCCCTGCTTAATTACC CCACAGTCGGCAGGCTCCCGGCGACACTGCGATCGATCGATGACGCC  
GCTGATGCCGCTCCCGTGTGATCGGCACAGGACGGTGTGGTCGGAACCGCCGAAGCGGCCGGCCGCTCCGACCAA  
GTTCAAGGAGACGCGCCACCCGCTGTACCGCGCGTGC GGCGACGGGGCCCCCGCGCGCGGTGGGTGTGCGAGGTGC  
GCGTGCTCGGGATGAGGGGCTCCAGGCTTTGGCTCGGCACCTTCACCACCGCCGAGATGGCAGCGCGCGCACGAC  
GCCGCCGTTCTCGCGCTCTCCGGCCGCGCGCTTGCTCAACTTCGCCGACTCCGCCTGGCGGATGCTCCCCGTCT  
CGCCGGCCCGTTTACGACGCGCAAGGAGATCAAGGATGCCGTGCGCGTCCGCGTCTTACGTTCCAAAGGCAGCACC  
CGTTCGCGTCCATGGCACCATTGTCCCTGCGCGGACAACCGATGACGAGAAGGAAATCGATGGCTTGCCGGCACCG  
AGCGCCCTGTCCATGTCCAGCGAGCTGTTGAATGAGCACTGGTTTGGCGGCATGGATGCCGGATCGTGCTACTCGGA  
GTTTCATGGAGTCGCCGGACACAGACCGTGGCGGGAAGACTTTGAGCTCGGTGGCGTCGAGACACCGCCATGGAGCT  
ACTTGTTTCGACTAAGCAGTTAAAGGTTTTTAGACGGTTTCTCTGCTTCTGTTCCA

>TaCBFIVd-4.1\_S90

ACACCATCGCTCAACGTGCTCAACCATCTCCACTAGCTCTCGACTCAGATGACGTCGCCGACGCTGCCTCCAAGTC  
CGCCAGCATGACAGGGTCACAGGACGCTGTGTCGGAGCGCGGAAGCGGCCGCGCGGGCGGACCAAGTTGCACG  
AGACGCGCCACCCGCTGTACCGGGCGTGC GGCGCGCGTGGCCGGGTCCGGCAGTGGGTGTGCGAGGTGCCGTGCC  
GGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGCACGACGCCGCGGT  
GCTCGCGCTCTCCGGCCGCGCGCTGCTCAACTTCGCCGACTCCGCATGGCGGATGCTGCCGTGCTTGCGGCCG  
AATCGCTTCGGCTTCGGCAGCGCGCGGAGATCAAGCTTGCCGTGCGCGTTCGCGTTCGCTTCCAGCAGCAGCAG  
ATTATTCTTCCAGTCGCGTGTCCAACGGTGGAGGCGCGCGCCAGCCGAGCAACTCTCTGTTTTACATGTCGTCCGT  
CGACTTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCTGGGTGCTACTACGAGAGCTTGCGCAGG  
GGATGCTCATGGCGCCGCCGACGACAGATGAGGCGGGAGGACGCCGAGCAGACCGGCGTCGAGACACCGACGCCG  
TTATGGAGCTATTTGTTGACTAATTCAGCACGCAGTGTAAAGTTGTCGATAGTTGCGTTGTGTTCTTCCAAAGGG  
GCAATTCAGCACACTGGCGCGCGTTACTAGTGGATCCGAGCTTCGGT

>TaCBFIVd-4.1\_S91 (manque\_39b\_pour\_ATG)

GACCAAGGTCACAGGACGCTGTGTCGGAGCGCGCGGAAGCGGCCGGCGGGCGGACCAAGTTGCACGAGACGCGCCA  
CCCGCTGTACCGGGCGTGC GGCGCGCGTGGCCGGGTCCGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGGTGAAGG  
GCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGCACGACGCCGCGGTGCTCGCGCTC  
TCCGGCCGCGCGCGCTGCCTCAACTTCGCCGACTCCGCATGGCGGATGCTGCCCGTGCTTGCGGCCGAATCGTTCCG  
CTTCGGCAGCGCGCGGAGATCAAGCTTGCCGTGCGCGTTCGCGTTCGCTTCCAGCAGCAGCAGATTATTCTTC  
CAGTCGCGTGTCCAACGGTGGAGGCGCGCGCCAGCCCGAGCAACTCTCTGTTTTACATGTCGTCCGTGCACTTGCTG  
GAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCTGGGTGCTACTACGAGAGCTTGCGCAGGGGATGCTCAT  
GGCGCCGCGGACGACAGAGCGAGGCGGGAGGACGCCGAGCAGACCGGCGTCGAGACACCGACGCCGTTATGGAGCT  
ATTTGTTTGACTAATTCAGCACGCAGTGTAAAGTTGTCGATAGTTGCGTTGTGTTCTTCCCA

>TaCBFIVd-4.1\_M124 (manque\_7b\_pour\_TAA)

ACACCATCGCTCAACGTGCTCAACCATCTCCACTAGCTCTCGACTCAGATGACGTCGCCGACGCTGCCTCCAAGTC  
CGCCAGCATGACAGGGTCACAGGACGCTGTGTCGGAGCGCGGAAGCGGCCGGCGGGCGGACCAAGTTCCACG



AGACGCGCCACCCGCTGTACCGGGGCGTGC GGCGCCGTGGCCGGGTGCGGCAGTGGGTGTGCGAGGTGCGCGTG  
 CCCGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGCACGACGCCG  
 GGTGCTCGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATGGCGGATGCTGCCCGTGTTCGCG  
 CC**GAA**TCGTTTCGGCTTCGGCAGCGCGCGGAGATCAAGCTTGCCGTCGCCGTTGCCGTCGTTTCCAGCAGCAG  
 CAGATTATTCTTCCAGTCGCGTGTCCAACGGTGGAGGCGGCCAGCCGAGCAACTCTCTGTTTTACATGTCGTC  
 CGTCGACTTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCTGGGTCTACTACGAGAGCTTGGCGC  
 AGGGGATGCTCATGGCGCCGCCGGACGACAGAGCGAGGCGGGAGGACGCCGAGCAGACCGGCGTCGAGACACCGACG  
 CCGTTATGGAGCTATTGTT

>TaCBFIVd-4.1\_M125 (manque 15b pour ATG)

GCTGCCTCCAACCCCCACCATCAGCAGGGTCAACAGGACGGTGTGTCGTCGGAGCCGCCGAAGCGGCCGGCGGGGCG  
 GACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGGGGCGTGC GGCGCCGTGGCCGGGTGCGGCAGTGGGTGTGCG  
 AGGTGCGCGTGC CGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGC  
 CACGACGCCCGGTGCTCGCGCTCTCCGGCCGCGCGCCCTGCCTCAACTTCGCCGACTCCGCATGGT**TGG**ATGCTGCC  
 CGTGCTTGCGGCC**GAA**TCGTTTCGGCTTCGGCAGCGCGCGGGAGATCAAGCTTGCCGTCGCCGTTGCCGTCGTTGCGT  
 TCCAGCAGCAGCAGATTATTCTTCCAGTCGCGTGTCCAACGGTGGAGGCGGCCAGCCGAGCAACTCTCTGTTT  
 TACATGTCGTCGTCGACTTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCTGGGTCTACTACGA  
 GAGCTTGGCGCAGGGGATGCTCATGGCGCCGCCGACGACAGAGCGAGGCGGGAGGACGCCGAGCAGACCGCGCTCG  
 AGACACCGACGCCGTTATGGAGCTATTTGTTTGA**TAA**TCAGCACGCAGTGTAAAGT**GTCGATAGTTGCGTTGTG**  
**TTCTTCCCA**

>TaCBFIVd-4.1\_N122

ACACCATCGCTCAACGTGCTCAACCATCTCCACTAGCTCTCGACTCAG**ATG**GACGTGCGCGACGCTGCCTCCAAGTC  
 CGGCCAGCATGAGCAGGGTCAACAGGACGGTGTGTCGTCGGAGCCGCCGA**ACCCG**CGCGGGGCGGACCAAGTTCCACG  
 AGACGCGCCACCCGCTGTACCGGGGCGTGC GGCGCCGTGGCCGGGTGCGGCAGTGGGTGTGCGAGGTGCGCGTGCCC  
 GGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGCACGACGCCCGGT  
 GCTCGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATGGCGGATGCTGCCCGTGTTCGCGCCG  
 GATCGTTTCGGCTTCGGCAGCGCGCGGGTGATCAAGCTTGCCGTCGCCGTTGCCGTCGTTGCGTTCCAGCAGCAGCAG  
 ATTATTCTTCCAGTCGCGTGTCCAACGGTGGAGGCGGCCGCCAGCCGAGCAACTCTCTGTTTTACATGTCGTCGCT  
 CGACTTGCTGGAGCTCGACGAGGAGCAGTGGTTTGGCGGCATGGACGCTGGGTCTACTACGAGAGCTTGGCGCAGG  
 GGATGCTCATGGCGCCGCCGACGACAGAGCGAGGCGGGAGGACGCCGAGCAGACCGGCGTCGAGACACCGACGCCG  
 TTATGGAGCTATTTGTTTGA**TAA**TCAGCACGCAGTGTAAAGT**GTCGATAGTTGCGTTGTGTTCTTCCCA**

>TaCBFIVd-B4\_S266

TCAG**ATG**GACGTGCGCGACGCTGCCTCCAAGTCCGGCCACCAGGAGCAGGGTCAACAGGACGGTGTCTCGGAGCCGC  
 CGAAGCGGCCCGCGCGCGGACCAAGTTCCGTGAGACACGACACCCGCTGTACCGCGCGTGC GGCGCCGTGGCCGG  
 GTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCAGGAGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGC  
 CGAGATGGCGGCGCGCGCACGACGCCGCGGTGCTCGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACT  
 CCGCATGGCGGATGCTGCCCGTGTTCGCGGCCGGTCTGTT**GCG**CTTCGGCAGCGCGCGGGAGATCAAGCTTGCCG**TC**  
**AG**CCGTCGCCGTCGTCGCTTCCAGCTGCAGCAGAAATATCTCCAGCCGCGTGTCCAACGGTGGAGCCGACCGACA  
 GCCCGAGCAACTCTCTGTTCTACATGTCGTCGGCGACCTGCTGGAGCTCGACGAGGAGCACTGGTTTGGCGGCATG  
 GATGCCGGGTCTGACTACGACAGCTTGGCGCAGGGATGCTCGTGGAGCCGCCGACGACAGAGCGAGGCGGGAGGA  
 CGCCGAGCAGACGGGCTCGAGACACCGACGCGCGCTATGGAGCTATTTGTTTGA**TAA**TCAGCACGCAGTGTAAAG  
 TTGTCGATA**GTTCTGTTGTGTTCTTCCCGATCTGG**

>TaCBFIVd-B4\_S267 (manque 17b pour ATG)

TGCCTCCAAGTCCCGCCACCGAGGACAGGGTCAACAGGACGGTGTCTCGGAGCCGCCGAAGCGGCCGGCGGGGCGGA  
 CCAAGTTCCGTGAGACACGACACCCGCTGTACCGCGCGTGC GGCGCCGTGGCCGGGTGCGGCAGTGGGTGTGCGAG  
 GTGCGCGTGCCAGGAGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGCA  
 CGACGCCGCGGTGCTCGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATGGCGGATGCTGCCCG  
 TGCTCGCGGCGGGTGTTC**GCG**CTTCGGCAGCGCGCGGGAGATCAAGCTTGCCG**TCAG**CCGTCGCCGTCGTCGCTT  
 CCAGCTGCAGCAGAAATATCTCCAGCCGCGTGTCCAACGGTGGAGCCGACCGACAGCCCGAGCAACTCTCTGTTCT



ACATGTCGTCGGCGACCTGCTGGAGCTCGACGAGGAGCACTGGTTTGGCGGCATGGATGCCGGGTCTGTACTA  
GACAGCTTGGCGCAGGGGATGCTCGTGGAGCCCGGACGACAGAGCGAGGCGGGAGGACGCCGAGCAGACGGGCT  
CGAGACACCGACGCCGCTACGGAGCTATTTGTTTGAATAATTCAGCACGCAGTGTAAGTTGTTCGATAAGTTCTGTG  
TGTCTTCCCGATCTGG

>TaCBFIVd-26.2\_S165

CAAACCTCTCTACACAGCTGCTGATTCTTCCAGTGCTCCACACTTGAGAAGCAGCAGAGATCTCCGCAAGCTCTTGA  
CCTAGATGCCGTCTGGTCAGGAGGAGCAGCGGCACAGGACGGTGAGGTCGGAGCCACCGAAGCGGCCCGCGGGCGG  
ACCAAGTTCCAGGAGACGCGCCACCCGCTGTACCGCGCGTGCGCCAGCGGGGGCCCGCGGGCGGTGGGTCTGCGA  
GGTGCGCGTGTCTGGGATGAGGGGCTCCAGGCTCTGGCTCGGCACCTTCGTACCGCCGAGATGGCGGCGCGCGCC  
ACGACGCGCGCTACTCGCGCTCTCCGGCCGAAAGCCTGCCCTCAACTTCGCCGACTCCGCCTGGCGGATGCTGCC  
GTGCTCGCGCGCGGCTCCTTCGGCTTCGGCAGCGCGCGGAGATCAAGACCGCGTCCGCGTCCGCGTCCGCGT  
CCAGCGGCAGCATCGTTCTTCCAGTCGCGTGCCCGCGCGGAGCCGCGCTCGCCCGAGCGGCGCGCTGTTCT  
CCATGTGCTCAGGCGACTTGCTGGAGCTCGACGACGAGCAGTGGTTTGGCGGCATGGTCGCCGGGTCTGTACTACGAG  
AGCTTGGCGCAGGGGATGCTCGTGGAGCCCGGACGCGCGGAGCTGGCGAGAGGACAGCGACCACAGCGCGTGGC  
GGAGACGACACTCCGTTGTGGAGCTAAATCGCTGAGCTAAGCAGACCAAAGTTTTACTGTACATATTAGTTTCGAA  
GAAACAGAGTAGGCAGTTTGTCTTCTCGTTGTGTATAGCA

>TaCBFIVd-26.2\_S166

CTGAGAAGCAGCAGAGATCTCCGCAAGCTCTTGACCTAGATGCAGTCTGGTCAGGAGGAGCAGCGGCACAGGACGGT  
GAGGTCTGGAGCCACCGAAGCGGCCCGCGGGCGGACCAAGTTCCAGGAGACGCGCCACCCGCTGTACCGCGCGTGC  
GCCAGCGGGGGCCCGCGGGCGGTGGGTCTGCGAGGTGCGCGTGTCTCGGGATGAGGGGCTCCAGGCTCTGGCTCGGC  
ACCTTCGTACCGCCGAGATGGCGGCGCGCGCCACGACGCCCGCTACTCGCGCTCTCCGGCCGAAAGCCTGCCT  
CAACTTCGCCGACTCCGCCTGGCGGATGCTGCCCGTGTCTCGCGCGCGGCTCCTTCGGCTTCGGCAGCGCGCGGAGA  
TCAAGACCGCGTCCGCGTCCGCGTCTCGCGTTCAGCGGCAGCAGATCGTTCTTCCAGTCGCGTGCCCGCGCGCG  
GAGCGGCGCGTCCGCCGAGCGGCGCGTGTCTCCATGTCTCAGGCGACTTGCTGGAGCTCGACGACGAGCAGTG  
GTTTGGCGGCATGGTCGCCGGGTCTGTACTACGAGAGCTTGGCGCAGGGGATGCTCGTGGAGCCCGGACGCCGAG  
CCTGGCGAGAGGACAGCGACCACAGCGCGTGGCGGAGACGACGCGGTTGTGGAGCTAAATCGCTGAGCTAAGCA  
GACCAAAGTTTTACTGTACATATTAGTTTCGGAAGAAACAGAGTAGGCAGTTTGTCTTCTCGTTGTGTATAGCA

>TaCBFIVd-26.2\_N140

AGTGCTCCACACTTGAGAAGCAGCAGAGATCTCCGCAAGCTCTTGACCTAGATGCCGTCTGGTCAGGAGGAGCAGCG  
GCACAGGACGGTGAGGTCTGGAGCCACCGAAGCGGCCCGCGGGCGGACCAAGTTCCAGGAGACGCGCCACCCGCTGT  
ACCGCGCGGTGCGCCAGCGGGGGCCCGCGGGCGGTGGGTCTGCGAGGTGCGCGTGTCTCGGGATGAGGGGCTCCAGG  
CTCTGGCTCGGCACCTTCGTACCGACGAGATGGCGGCGCGCGCCACGACGCGCGCGTACTCGCGCTCTCCGGCCG  
CAAGCCTGCCTCAACTTCGCCGACTCCGCTGGCGGATGCTGCCCGTGTCTCGCGCGCGGCTCCTTCGGCTTCGGCA  
GCGCGCGGAGATCAAGACCGCGTCCGCGTCCGCGTCTCGCGTTCAGCGGCAGCAGATCGTTCTTCCAGTCGCG  
TCCCGCGCGGAGCCGCGCGTCCGCGCGAGCGGCGCGCTGTTCTCCATGTCTGTCAGGCGACTTGCTGGAGCTCGA  
CGACGAGCAGTGGTTTGGCGGCATGGTCGCCGGGTCTGTACTACGAGAGCTTGGCGCAGGGGATGCTCGTGGAGCCG  
GCGCGCGGAGCTTGGCGAGAGGACAGCGACCACAGCGCGTGGCGGAGACGACGCGGTTGTGGAGCTAAATCG  
CTGAGCTAAGCAGACCAAAGTTTACTGTACATATTAGTTTCGGAAGAAACAGAGTAGGCAGTTTGTCTTCTCGTTG  
GTATAGCA

>TaCBFIVd-9.1b\_S128 (manque\_88b\_pour\_ATG)

GGCGGACCAACTTCCACGAGACGCGCCACCGCTGTACCGCGGCGTGGCGCGCGTGGCCGCTCGGGCAGTGGGTG  
TGCGAGGTGCGCGTGGCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATGGCGGCGCG  
CGCGCACGACGCGCGCTGCTCGCGCTCTCCGGCCGCGCGCTGCCTCAACTTCGCCGACTCCGATGGCGCATGC  
TGCCCGTGTCTCGCGCCGGCTCCTTCGGCTTCGGCAGCGCGAGCGAGATCAAGGCCCGCGTCTGCTGCGCGTCTG  
GCGTTCCTGCGGAAGCAGATTGTTCTTCCAGTCGCGCTCGCGTCTGCGGCTCCAGCAGAAGCAGGTCCCGATCGC  
CGTCCGCGTGGTGGCGCTCCAGCAGAAGCAGGTTCCGCTCGCGTCCGCTGGTGGCGCTCCAGAAGCTGCCAGTTC  
CGGTGCGCGTCCGCTGCTGGCGCTCCAGCAGCAGCAGATCATCTTCCAGTCGCGTGCCTGGCGCCGAGTTTAC  
ATGCTTCCGGCGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGCCGGGTCTGTACTACGCGAG

CTTGGCGCAGGGGATGCTCGTGGCGCCGCCGGACGAAAGAGCCAGGCCGGAGAGCGGCGAGCAGAGCGGGCTCC  
AGACGCCGCTGTGGAGCTGCTTGTTCGACTAAATTAGCGATACTGTCAAGTTGTAGATAGTCGCCT

>TaCBFIVd-9.1b\_S131 (manque\_138b\_pour\_ATG)

CGCGTGGCCGCGTTCGGGCAGTGGGTGTGCGAGGTCCCGTCCCGGGATCAAGGGCTCCAGGCTCTGGCTCGGCAC  
CTTCAACACGGCCGAGATGGCGGCGCGCGCATGACGCCGCCGTGCTCGCGCTCTCCGGCCGCGCCGCTGCCTCA  
ACTTCGCCGACTCCGCATGGCGCATGCTGCCGTGCTCGCGGCCGGCTCCTTCGGCTTCGGCAGCGCGAGCGAGATC  
AAGCCGCCGCTCGCTGTGCGCGTCTGCGGTTCTTGCAGGAGCAGATTGTTCTTCAGTCGCCGTCGCCGTCGTGGC  
GCTCCAGCAGAAGCAGGTCCCGATCGCCGTCGCCGTGGTGGCGCTCCAGCAGAAGCAGGTTCCGGTCGCCGTCGCCG  
TGGTGGCGCTCCAGAAGCTGCCAGTTCGGTTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATCATTCTTCCA  
GTCGCGTGCCTGGCGCCCGAGTTTTACATGTCTTCCGGCGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGG  
CATGGAGGCCGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGGACGAAAGAGCCAGGCCGG  
AGAGCGGCGAGCAGCGCGTCCAGACGCCGCTATGGAGCTGCTTGTTCGACTAAATTAGCGATACTGTCAAGTTG  
TAGATAGTCGCCT

>TaCBFIVd-9.1b\_M130 (manque\_174b\_pour\_TAA)

CAACACAGCCGTTGATTCTTCCACTACTCCCGCTCCACACCTCCACGAGCGTCTCCGCCAGCTCTCGACTCAGATG  
GACGTTCGCCGACATCCCTCCCGCTCCCGCCACACAGCAGGGCCACCGACGGTGTGCTCGGAGCCGCCGAGAGG  
GCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCAGCGCCGTGGCCGCGTCCGGC  
AGTGGGTGTGCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATG  
GCGGCGCGCGCGCACGACGCCCGCGTGTGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATG  
GCGCATGCTGCCGTGCTGCGGCGCGGCTCCTTCGGCTTCGGCAGCGCGAGCGAGATCAAGGCCGCCGTCGTGTGCG  
CCGTGCTGCGGTTCCCGCGGAAGCAGATTGTTCTTCCAGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGAAGCAGGTC  
CCGATCGCCGTCGCCGTGGTGGCGCTCCAGCAGAAGCAGGTTCCGGTCGCCGTCGCCGTGGTGGCGCTCCAGAAGCT  
GCCAGTTCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATCATTCTTCCAGTCGCGTGCCTGGCGCCCG  
AGTTTTACATGTCTTCC

>TaCBFIVd-9.1a\_N141

CAACACAGCCGTTGATTCTTCCACTACTCCCGCTCCACACCTCCACGAGCGTCTCCGCCAGCTCTCGACTCAGATG  
GACGTTCGCCGACATCCCTCCCGCTCCCGCCACACAGCAGGGCCACCGACGGTGTGCTCGGAGCCGCCGAGAGG  
GCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCAGCGCCGTGGCCGCGTCCGGC  
AGTGGGTGTGCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATG  
GCGGCGCGCGCGCACGACGCCCGCGTGTGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCCGCTG  
GCGCATGCTGCCGTGCTGCGGCGCGGCTCCTTCGGCTTCGGCAGCGCGAGCGAGATCAAGGCCGCCGTCGTGTGCG  
CCGTGGTTCGCTTCCAGCGGAAGCAGATTGTTCTTCCAGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGAAGCAGGTC  
CCGATCGCCGTCGCCGTGGTGGCGCTCCGGCAGAAGCAGGTTCCGGTCGCCGTCGCCGTGGTGGCGCTCCAGCAGCT  
GCCAGTTCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATCATTCTTCCAGTCGCGTGCCTGGCGCCCG  
AGTTTTACATGTGCTCCGGCGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGTCTGAC  
TACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGGACGAAAGAGCGAGGCCGGAGAGCGGCGAGCAGAGCGG  
CGTCCAGACGCCGCTATGGAGCTGCTTGTTCGACTAAATTAGCGA

>TaCBFIVd-B9\_S132 (manque\_62b\_pour\_TAA)

GCCGCTGATTCCATCACTACTACCACTCCACACCTCTACGAGCATCTCCGCTAGCTCTCGACTCAGATGGACGTCG  
CCGACATCGCTCCCGCTCCCGCCAGCAGGAGCAGGGCCACCGAACGGTGTGCTCGGAGCCGCCGAGAGGCTGCG  
GGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCAGCGCCGTGGCCGCGTCCGGCAGTGGGT  
GTGCGAGGTGCGCGTGCCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATGGCGCGC  
GCGCGCACGACGCCCGGTGCTGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCCTGGCGGATG  
CTGCCCGTGTGCTGCGGCGCGGCTCGTTCGGCTTTGACAGCGCGCGGGAGATCAAGGCCGCCGTCGCCGTGGCGCTG  
CGCGTTCCAGCGGAAGCAGATCATTCCAGTCGCCGTCGCCGTCGTTGCGCTCCAGCAACAGCAGGTTCCGGTTCGAG  
TGGCGGTGGTGGCGCTCCAGCAGAAGCAGGTTCCCGCTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCTGCAAGTTCCG  
GTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATCATTCTTCCAGTCGCGTGCCTGGCGCCGAGTTTAA



CATGTCTTCCGGCGACCTGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGGCCGGGTCTACTACG  
CGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGA

>TaCBFIVd-B9\_S134 (manque\_70b\_pour\_ATG)

C6CCGAAGCGCCCTCCGCGCCCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGC GGCGCCGTGGC  
C6CGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACAC  
GGCCGAGATGGCGGCGCGCGCACGACGCCGCGGTGCTCGCGCTCTCCGGCCGCGCCGCTGCCTCAACTTCGCCG  
ACTCCGCTTGGCGGATGCTGCCCCGTGCTCGCGGCCGGCTCGTTTCGGCTTTGACAGCGCGCGGGAGATCAAGGCCGCC  
GTCGCCGTGGCCGTCTGTCGCTTCCAGCGGAAGCAGATCATTCCAGTCGCCGTGCGCGTCTGTCGCTCCAGCAACA  
GCAGGTTCCGGTTCGAGTGGCCGTGGTGGCGCTCCAGCAGAAGCAGGTCCCGCTCGCCGTGCGCGTCTGTCGCGCTCC  
AGCAGCTGCAGGTTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCAGCAGATCATTCTTCCAGTCGCGTGC  
CTGGCGCCGAGTTTTACATGTCTTCCGGCGACCCGTTGGAGCTGGACGAGGAGCAGTGGTTTGGCGGCATGGAGGC  
CGGTTCGTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGCGAGCAGCGCG  
AGCAGAGCGCGCTCCAGACGCCACTATGGAGCTGCTTGTTCGACTAATTAGCACATTGCCAAGTTGTAGATAGTC  
GCGCTCTT

>TaCBFIVd-D9\_S135 (manque\_17b\_pour\_TAA)

CCGCTGATTCTTCCAGTACTCTGTCTCCACACCTCCCACGAGCGTCTCCGCCAGCTCTCGACTCAGATGGACGTGCG  
CGACATCGCCTCCCGCTCTGCGCAGCAGGGGCATCGGGGACACCGGACGGTGTCTGTCGGAGCCGCGCGAGCGCCCG  
CGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGC GGCGCCGTGGCCGCGTTCGGGCAGTGG  
GTGTGCGAGGTGCGCGTGCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATGGCGGC  
GC GCGCGCACGACGCCGTGTGCTCGCGCTCATCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATGGCGCA  
TGCTGCCCGTGTCTCGCGCCGGGTCTTTCGGTTTCGGCAGCGCGAGCGAGATCAAGGCCGCGCTCGCCGTGGCCGT  
GTCGCGTTCCAGCGGAAGCAGATTATTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCAGCAGGTTCCGGTTCGC  
CGTCGCCGTCTGTCGCGCTCCAGCAGAAGCAGGTTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCTGCAGGTT  
CGGTGCGCGTGGCCGTCTGTCGCGCTCCAGCAGCAGCAGATTATTCTTCCAGTCGCGTGCCTGGCGCCGAGTTTTAC  
ATGTCTTCCGGCGACCTGTTGGAGCTTGATGAGGAGCAGTGGTTTGGCGGCATGGACGTGGGTCTGCTACTACGCGAG  
CTTGGCGCAGGGGATGCTCGTGGCGCCGCCGACGAAAGAGCGAGGCCGAGCACGGCGAGCAGAGCGGCGTCCAGA  
CGCCGCTATGGA

>TaCBFIVd-D9\_S274 (manque\_172b\_pour\_TAA)

CCGCTGATTCTTCCAGTACTCTGTCTCCACACCTCCCACGAGCGTCTCCGCCAGCTCTCGACTCAGATGGACGTGCG  
CGACATCGCCTCCCGCTCTGCGCAGCAGGGGATATCTCGGGGACACCGGACGGTGTCTGTCGGAGCCGCGCGAGCGCCCG  
CGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGTGC GGCGCCGTGGCCGCGTTCGGGCAGTGG  
GTGTGCGAGGTGCGCGTGCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATGGCGGC  
GCGCGCGCACGACGCCGTGTGCTCGCGCTCATCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATGGCGCA  
TGCTGCCCGTGTCTCGCGCCGGGTCTTTCGGTTTCGGCAGCGCGAGCGAGATCAAGGCCGCGCTCGCCGTGGCCGT  
GTCGCGTTCCAGCGGAAGCAGATTATTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCAGCAGGTTCCGGTTCGC  
CGTCGCCGTCTGTCGCGCTCCAGCAGAAGCAGGTTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCTGCAGGTT  
CGGTGCGCGTGGCCGTCTGTCGCGCTCCAGCAGCAGCAGATTATTCTTCCAGTCGCGTGCCTGGCGCCGAGTTTTAC  
ATGTCTTCCGG

>TaCBFIVd-D9\_S275 (manque\_153b\_pour\_TAA)

CCGCTGATTCTTCCAGTACTCTGTCTCCACACCTCCCACGAGCGTCTCCGCCAGCTCTCGACTCAGATGGACGTGCG  
CGACATCGCCTCCCGCTCTGCGCAGCAGGGGATATCTCGGGGACACCGGACGGTGTCTGTCGGAGCCGCGCGAGCGCCCG  
CGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGCGGAGCGGCGCCGTGGCCGCGTTCGGGCAGTGG  
GTGTGCGAGGTGCGCGTGCTGGGATCAAGGGCTCCAGGCTCTGGCTCGGCACCTTCAACACGGCCGAGATGGCGGC  
GCGCGCGCACGACGCCGTGTGCTCGCGCTCATCGGCCGCGCCGCTGCCTCAACTTCGCCGACTCCGCATGGCGCA  
TGCTGCCCGTGTCTCGCGCCGGGTCTTTCGGTTTCGGCAGCGCGAGCGAGATCAAGGCCGCGCTCGCCGTGGCCGT  
GTCGCGTTCCAGCGGAAGCAGATTATTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCAGCAGGTTCCGGTTCGC  
CGTCGCCGTCTGTCGCGCTCCAGCAGAAGCAGGTTCCGGTTCGCCGTGCGCGTCTGTCGCGCTCCAGCAGCTGCAGGTT  
CGGTGCGCGTGGCCGTCTGTCGCGCTCCAGCAGCAGCAGATTATTCTTCCAGTCGCGTGCCTGGCGCCGAGTTTTAC  
ATGTCTTCCGG



CGGTCGCCGTGGCCGTCGTGGCGCTCCAGCAGCAGCAGATTATTCTTCCAGTCGCGTGCCGTGGCGCCGGAGTTT  
TACATGTCCTTCCGGCGACCTGTTGGAGCTTGAC

>TaCBFIVd\_D9\_M136 (manque\_136b\_pour\_ATG)

TGCGGCGCCGTGGCCGCTCGGGCAGTGGGTCTCCGACCTGCGCGTGCCGTGGGATCAAGGGCTCCAGGCTCTGGCTC  
GGCACCTTCAACACGGCCGAGATGGCGGCGCGCGCACGACGCCGCTGTGCTCGCGCTCATCGGCCGCGCCGCTG  
CCTCAACTTCGCCGACTCCGCATGGCGCATGCTGCCCCTGCTCGCGGCCGGGTTCGTTTCGGTTTCGGCAGCGCGAGCG  
AGATCAAGGCCGCCGTCGCCGTGGCCGTCGTGCGGTTCCAGCGGAAGCAGATTATTCCGGTCGCCGTCGCCGTCGTG  
GCGCTCCAGCAGCAGCAGGTTCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGTAGAAGCAGGTTCCGGTCGCCGTCGGC  
CGTCGTGGCGCTCCAGCAGCTGCAGGTTCCGGTCGCCGTCGCCGTCGTGGCGCTCCAGCAGCAGCAGATTATTCTTC  
CAGTCGCGTGCCGTGGCGCCGGAGTTTTACATGTCCTTCCGGCGACCTGTTGGAGCTTGACGAGGAGCAGTGGTTTGGC  
GGCATGGACGCTGGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGCGCGCCGGACGAAAGAGCGAGGCC  
GGAGCACGGCGAGCAGAGCGGCGTCCAGACGCCGCTATGGAGCTGCTTGCTCGACTAATTAGAACTACTTCAAGT  
TGTAGATTTAGGAAGAAACATACTAGGCAG

>TaCBFIVd-B22\_N15

CCTCCAGTTCTCAGCACACATCGACCTGCTCGACCGTCTCCACTAGCGCTCGTCCAAGATGACGTCGCCGACGCG  
CCTCCCCGTCTGGCCAGGAGCAGCAGGGGCACAGGACGGTGTGCTCTGAGCCGCCGAGGCCAACC CGCGGGGCGGACG  
AAGGTACATGAGACGCGCCACCCGCTGTACCGCGCGTGCGGCAGCGGGGCCGGGTTCGGCAGTGGGTGTGCGAGGT  
GCGCGTGCCCGGGGTGAAGGGCTCCAGGCTCTGGCTCGGCACCTTACCAACGCCGAGATGGCGGCGCGCGCAGC  
ACGTCGCCGTGCTCGCGCTCTCCGGCCGGGTGCTGCTTCACTTCGCCGACTCCGCCTGGCGCATGCTGCCCGTG  
CTCGCGGCCGGGTCTTTCGGCTTCGGCAGCGCGCGGAGATCAAGGCCCGCTCGCCGTCGCCGTCGTGGCTTTCC  
GAAGGAGCAGATTATCCAGTCGCCGTCGCCGTGGTGCGATCCAGAAGCAGCAGATTATCCAGTCGCTGTGCTG  
TCGTGGCGATCCAGAAGCAGCAGATTATCCCGTCGCCATCGTGCGCTCCAGAAGCAGCAGATTCCAGTGGCGCT  
CCCTCGTGCGCTCCAGGAACAGCAGGTCCAGTCGCCGTCGCCGTCGTGGCGCTTCATCGGCAGCAGGTTCCAGT  
GCGTGCCCGGGAACCTCCGGCCCGGGCAGCGCTCTGTTTTACATGTCGTCCAGCGACTTGTGGAGCTCGACGAG  
GCAGTGGTTTGGCGGCATGGAGGCCGGTCTACTACGCGAGCTTGGCGCAGGGGATGCTCGTGCCACGCCCGGAT  
AAAGAGCGAGGACGGAGGACGGCGAGCAGAGCGGCGTCCAGACGCCGCTATGGAGCCACTTGTTCACCTAATCTAGC  
AGTGTA

Tableau S16. Amorces spécifiques utilisées pour ajouter des informations sur la séquence des gènes *TaCBF*.

Gène	Numéro d'accension	Amorce	Direction	Séquence de l'amorce
<i>TaCBFIIIc-3.1</i>	3.2_F1 <sup>a</sup>		Sens	5'-GCTCAACAAAGCTAGCTACACTCCTCAGT-3'
	3.1_R1 <sup>a</sup>		Anti-sens	5'-ATCAGCTGCCAACTGGGAATTATCGAC-3'
<i>TaCBFIVa-A2</i>	EF028769	C2_630_91_R2	Anti-sens	5'-AAGCTCAGGTGTGCCCTTCCGGGAGTAG-3'
<i>TaCBFIVa-2.2</i>	EF028770	IVa_2.2_R1	Anti-sens	5'-CCTTCCCCCGATGCTTACTCTCTGTTCTTCC-3'
<i>TaCBFIVa-2.3</i>	EF028771	IVa_2.3_F2	Sens	5'-GACGCCAGCTGCGACTGATGGACAC-3'
<i>TaCBFIVd-B22</i>	EF028786	IVd_A22_14B_R2	Anti-sens	5'-CCCCAAAGTAGAAACCGGGAATTGACAAGTGCTAC-3'

Les amorces sens spécifiques aux gènes ont été utilisées avec M13F et des amorces anti-sens spécifiques aux gènes ont été utilisées avec M13R pour amplifier les régions 3' ou 5' des ADNc de *TaCBF*, respectivement. Les séquences supplémentaires ont été utilisées pour mettre à jour les fichiers d'accension GenBank pour présenter la plus longue séquence.

<sup>a</sup>, Ces amorces ont été utilisées pour amplifier le gène à partir de l'ADN génomique.

Tableau S27. Amorces spécifiques utilisées pour amplifier les nouveaux gènes *TaCBF*.

Gène	Nom du clone	Amorce	Direction	Séquence de l'amorce
<i>TaCBFI-1.1</i>	C1_CA601_06_R2L1_562	C1_CA601_65_F2	Sens	5'-GCAACAGCGAGGGCTTCTCGACGTACA-3'
		C1_CA601_06_R2	Anti-sens	5'-CTGGATTGCAATATAAAGCTGCACGTCCATTAGAA-3'
<i>TaCBFI-1.2</i>	ContigCA601682R1_L6_65	C1_CA601_65_F2	Sens	5'-GCAACAGCGAGGGCTTCTCGACGTACA-3'
		C1_CA601_65_R1	Anti-sens	5'-TCTTCGCGCCCCCAAGGCATAAATC-3'
<i>TaCBFI-1.3</i>	ContigCA601682R2_L1_04	C1_CA601_65_F2	Sens	5'-GCAACAGCGAGGGCTTCTCGACGTACA-3'
		CA601682R2	Anti-sens	5'-CAGAAGAGAAAATGGAAGGCAACCATGCAG-3'
<i>TaCBFI-5.4</i>	C5_63098601F1_L8_124	5_63098601F1	Sens	5'-GCCAACACCAATTGATTGCTAGCTAGCTACAGTGTCT-3'
		5_63098601R2	Anti-sens	5'-TCATGCGAATATGGTGTGCGTCCATACG-3'
<i>TaCBFI-5.5</i>	C5A_DQ353853R2_L6_170	5A_DQ353853F2	Sens	5'-GCGGTGATACCTTGCAGAGATGGAACAGTG-3'
		C5A_DQ35_170_R2	Anti-sens	5'-TGCGCATCTTTCCTGACAAAGCTGTAGTGTG-3'
<i>TaCBFI-23.1</i>	C18_63098599R1_L6_31	C18_630_31_F1	Sens	5'-CACTCTCTCACTCCAGCGCAATCATCAAGC-3'
		63098599R1	Anti-sens	5'-CGGCGCAGTCTCCATTGTCCAGTACG-3'
<i>TaCBFI-23.2</i>	C18_63098599R1_L6_40	C18_630R1L6_40_F1	Sens	5'-CTCACTCCAGCGCAATCTCAAGCTAGACAC-3'
		63098599R1	Anti-sens	5'-CGGCGCAGTCTCCATTGTCCAGTACG-3'
<i>TaCBFI-23.3</i>	C18_630R1L6_40F1L1_822	C18_630R1L6_40_F1	Sens	5'-CTCACTCCAGCGCAATCTCAAGCTAGACAC-3'
		10_63098607F1	Sens	5'-TCGCACCGGAGTAGCTCACACTCTCACT-3'
<i>TaCBFI-10.1</i>	C10_63098607R1_L1_214	10_63098607R1	Anti-sens	5'-ACAGAGTCAAAATAGACAGATCGGCTACAAAGCTC-3'
		C10_630_214_F1	Sens	5'-GCGAACACTCCCACTCAAGCTCAAGTATCTCG-3'
<i>TaCBFI-10.2</i>	C10_630_214_F1L1_739	C10_630_214F1L1_739_R2	Anti-sens	5'-GGCCTACAGAGTCTCTCCCGAAATGGA-3'
		13_63098609F1	Sens	5'-CCACGGCTAAGCTCACACTCTCTCGATAAGC-3'
<i>TaCBFI-13.1</i>	C13_63098609R1_L1_203	13_63098609R1	Anti-sens	5'-GTCAAAACAGAGCAGAAATCAGATGGGAATCA-3'
		C16_630_185_F1	Sens	5'-AGCTCAACCTGATCTCACACTCCACAGACCAC-3'
<i>TaCBFI-16.1</i>	C16_63098611R2_L1_185	16_63098611R2	Anti-sens	5'-GCTTCAGAACCGCATCTGATGCTATATACCA-3'
		C16_630_190_F1	Sens	5'-AGCTCAACCTGCTCTCACACTCCACAGAACAA-3'
<i>TaCBFI-16.2</i>	C16_630_190_F2L1_817	16_63098611R2	Anti-sens	5'-GCTTCAGAACCGCATCTGATGCTATATACCA-3'
		C16_630_190_F2	Sens	5'-GACCACTGCGCGATGGACATGA-3'
<i>TaCBFI-16.3</i>	C17_63098615F1_L8_133	17_63098615F1	Anti-sens	5'-CAGCATCTGATGCTATGCTATGCCAAAGGATTCA-3'
		C17_630_133_R3	Sens	5'-GAGCAGCGAGTAGCTAGTACTCACCGTTCAC-3'
<i>TaCBFI-17.1</i>	C17_63098615F1_L8_133	C17_630_133_R3	Anti-sens	5'-GTGACACGTCAATGCCAAAGGCAACACAT-3'
		C17_630_133_R3	Sens	5'-AGTGGCAGTCGCGTCCAGCAAAAAG-3'
<i>TaCBFI-24.1</i>	C17_63098615F1_L8_133	C17_630_133_R3	Anti-sens	5'-CAGTGTGAACCCGTGGAGGTTATCGACAT-3'
		C17_630_133_R3	Sens	5'-CAGTGTGAACCCGTGGAGGTTATCGACAT-3'
<i>TaCBFI-24.2</i>	C17_63098615F1_L8_133	C17_630_133_R3	Anti-sens	5'-CAGTGTGAACCCGTGGAGGTTATCGACAT-3'
		C17_630_133_R3	Sens	5'-AGTGGCAATCGCGTCCAGCAAAAAG-3'
<i>TaCBFI-24.3</i>	C2_63098615F2_L1_91	C2_630_91_R1	Anti-sens	5'-ACCCGCTCTCGACGAGTATGGACAC-3'
		C2_630_91_R1	Sens	5'-GAAACCGAGAGGTCTCACATTGAGGAAACCTT-3'



<i>TaCBFIVa-25.2</i>	C2_63098615F2_L1_93	2_63098615F2 C2_630_93R2	Sens Anti-sens	5'-ACCGCTCTCGACGAGTGATGGACAC-3' 5'-CATGTCCATGGAGCGAATGATTGATCTAGTCTAGTG-3'
<i>TaCBFIVb-21.2</i>	CIVb_21.1_12A_FIL1_611	IVb_21_1_12A_F1 21_1_12A_F1_L1_611_R2	Sens Anti-sens	5'-CCCGTTGCCAGTAGCCCTAACAACCTCTTC-3' 5'-CCCATATTGGTGGACAGAAAGTACAGAACCAACT-3'
<i>TaCBFIVd-26.1</i>	C4_63098615R1_L1_226	4_63098615F1 4_63098615R1	Sens Anti-sens	5'-CAAAACCTCCACCGATCCATCGATCAAAACC-3' 5'-GTCTGTAGCTGAACGATTAGCTCCACAACG-3'
<i>TaCBFIVd-26.2</i>	C4_63098615R2_L6_245	4_63098615F1 4_63098615R2	Sens Anti-sens	5'-CAAAACCTCCACCGATCCATCGATCAAAACC-3' 5'-GCCCTTCTTCCAAAAATAGAAAGCGGAAACTTGC-3'

Les amorces sens spécifiques aux gènes ont été utilisées avec M13F et des amorces anti-sens spécifiques aux gènes ont été utilisées avec M13R pour amplifier les régions 3' ou 5' des ADNc de TaCBF, respectivement. Lorsque les deux régions ont été amplifiées, les sections qui se chevauchent (100% d'identité) ont été fusionnées pour présenter la plus longue séquence.

Tableau S38. Amorces spécifiques utilisées pour amplifier les gènes *TaCBF* à partir de l'ADN génomique de différents cultivars de blé.

Gène	Amorce	Direction	Séquence de l'amorce
<i>TaCBFI-1.1</i>	1.1_F1	Sens	5'-TTCCACTCCCTCCCCAGCTTACA-3'
	1.1_R1	Anti-sens	5'-GGAGTTCGAACACTCACAAGCCATAACC-3'
<i>TaCBFI-1.2</i>	1.2_F1	Sens	5'-TTCCACTCCACCCAGCTTACA-3'
	1.2_R1	Anti-sens	5'-CGGAGTTCGAACACTCACAATCCATAACC-3'
<i>TaCBFI-1.3</i>	1.3_F1	Sens	5'-CTTCCACTTCTCCCCAGCTTACAC-3'
	1.3_R1	Anti-sens	5'-CCTTCGGAGTTAAACACTAACAGCCATAACC-3'
<i>TaCBFI-5.1</i>	5.1_F1	Sens	5'-AAGGCTACCCACACAAACACCAC-3'
	5.1_R1	Anti-sens	5'-GAGCAAGCCAAATCGGGCTAGAGAC-3'
<i>TaCBFI-5.2</i>	5.2_F1	Sens	5'-CAACATCACCTCACTCACCAGTCACCA-3'
	5.2_R1	Anti-sens	5'-GGAGCAAGCCAAATCGAGATGAAGAATG-3'
<i>TaCBFI-5.3</i>	5.3_F1	Sens	5'-TCACCTCACCTCACTCAACCACCATG-3'
	5.3_R1	Anti-sens	5'-GAAAGCTGACCGGGCTGGAGATG-3'
<i>TaCBFI-5.4</i>	5.4_F1	Sens	5'-CCAACACCATTTGCTAGCTACAGT-3'
	5.4_R1	Anti-sens	5'-CGCTCAGCAAACTGGCTGCATATATTC-3'
<i>TaCBFI-5.5</i>	5.5_F1	Sens	5'-CAACACCATTTGCTAGCTACGGACA-3'
	5.5_R1	Anti-sens	5'-AACTGGTTGCGCATCTTCTCTGAC-3'
<i>TaCBFI-6.1</i>	6.1_F1	Sens	5'-GCCAAGGAACAGAGCATCGAAGGAA-3'
	6.1_R1	Anti-sens	5'-GTAAATATTCGAACTGGCGCAACTAGC-3'
<i>TaCBFI-6.2</i>	6.2_F1	Sens	5'-CCAAAAGCATCTCAGCAACAAGAACCC-3'
	6.2_R1	Anti-sens	5'-AGCTAAAACCTCGAAGCAGCGGAACTA-3'
<i>TaCBFI-6.6</i>	D6_F1	Sens	5'-AAAACCACTTCACTCGAGCAACACC-3'
	D6_R1	Anti-sens	5'-GGTAAATATTTCTAACTGGCGCACTATTGC-3'
<i>TaCBFI-23.1</i>	23.1_F1	Sens	5'-TCTCACTCCAGCGCAATCATCAAGC-3'
	23.1_R1	Anti-sens	5'-TGAGAGGAATGAGGAGACACGCAGAAC-3'
<i>TaCBFI-23.2</i>	23.1_F1	Sens	5'-TCTCACTCCAGCGCAATCATCAAGC-3'
	23.2_F1	Sens	5'-ACTAGCTATCGACCAAGCAGCCAGCAG-3'
<i>TaCBFI-23.3</i>	23.1_R2	Anti-sens	5'-GGACTGTCGGGACGTGGATCAAA-3'
	23.1_F1	Sens	5'-TCTCACTCCAGCGCAATCATCAAGC-3'
<i>TaCBFI-3.1</i>	23.1_R1	Anti-sens	5'-TGAGAGGAATGAGGAGACACGCAGAAC-3'
	3.2_F1	Sens	5'-GCTCAACAAGTAGCTCACACTCTCAGT-3'
<i>TaCBFI-3.2</i>	3.1_R1	Anti-sens	5'-ATCAGCTGCCAACTGGGAATTATCGAC-3'
	3.2_F1	Sens	5'-GCTCAACAAGTAGCTCACACTCTCAGT-3'
<i>TaCBFI-3.3</i>	3.2_R1	Anti-sens	5'-TTTGTGGAGGGGAATTATCGACTGTACTA-3'
	D3_F1	Sens	5'-ACAATTAGCTCAGGGCCTCACACTCCT-3'
<i>TaCBFI-B10</i>	D3_R1	Anti-sens	5'-TTTGTGGGGGATTATCGACACTA-3'
	B10_F1	Sens	5'-CAAGGAAGCCACCTGCCAATCAC-3'
	B10_R1	Anti-sens	5'-CGCTACAGAGTCAAAACAAAGCAGAAATCG-3'



<i>TaCBFIIIc-10.1</i>	10.1_F1 10.1_R1	Sens Anti-sens	5'-ACCAAGCTCGACTGCTCAGAAAAGGAAG-3' 5'-CGTCTCAGCTAGAGAAGAACGGAGTATAAACAG-3'
<i>TaCBFIIIc-10.2</i>	10.2_F1 10.2_R1	Sens Anti-sens	5'-GCACCAAGCTCGACTGCTCCATAAG-3' 5'-CGAAATTGGAAAACITTTGCGAAGGAAA-3'
<i>TaCBFIIIc-13.1</i>	13.1_F1 13.1_R1	Sens Anti-sens	5'-GATCACCCGACGCTCTCAATCAAG-3' 5'-AGAATCAGATGGGAATCATTTATCGACAC-3'
<i>TaCBFIIIc-12.1</i>	12.1_F1 12.1_R1	Sens Anti-sens	5'-TACTCCAGCCAGCCAACTAGCCTAGCA-3' 5'-CCTATCACTGGCCTTTTATTATGCCCATC-3'
<i>TaCBFIIIc-B12</i>	B12_F1 B12_R1	Sens Anti-sens	5'-GCTCCTCCTCCAGTCAACTAGTCAAGCA-3' 5'-TCCCGGATTTTATTATGCCGATGAT-3'
<i>TaCBFIIIc-A15</i>	A15_F1 A15_R1	Sens Anti-sens	5'-CCATAGATCGACCTTGGCTCACCGTA-3' 5'-CAGGATTTGATGGGTTCAGTTCCACA-3'
<i>TaCBFIIIc-15.2</i>	15.2_F1 15.2_R1	Sens Anti-sens	5'-CGACCACTGCAGTACCAATGGA-3' 5'-TCAAAGCAGCTGGCTGGAGTGTTTTAGTAG-3'
<i>TaCBFIIIc-16.1</i>	16.1_F1 16.1_R1	Sens Anti-sens	5'-CAACCTGATCTCACACTCCACAGACCA-3' 5'-CGATAGAGTCAAAAGCAGCTGGTTTCAGCA-3'
<i>TaCBFIIIc-16.2</i>	16.2_F1 16.2_R1	Sens Anti-sens	5'-CCTGCTCTCACACTCCACAGAACACC-3' 5'-TGGAAATCAACTGCCATAGAGTGTCAAAGC-3'
<i>TaCBFIIIc-A19</i>	A19_F1 A19_R1	Sens Anti-sens	5'-GCTCAAGCTCAACTAGAGTGACCAATTCC-3' 5'-TGAAACCCACTGGAGTTACAGACATCAG-3'
<i>TaCBFIIIc-B19</i>	B19_F1 B19_R1	Sens Anti-sens	5'-CAAGCTCAGTAGACCGACCAATTCC-3' 5'-CAGTGTGAACCCACTGAAGTTACTGACATC-3'
<i>TaCBFIIIc-D19</i>	D19_F1 D19_R1	Sens Anti-sens	5'-GCGATCGGCACTCCGACAGCTA-3' 5'-GTGTGAACCCACTGCAGGTTACAGACA-3'
<i>TaCBFIIIc-24.2</i>	24.2_F1 24.1_R1	Sens Anti-sens	5'-CTCCACACGTACCCAAAGTGGCAATC-3' 5'-CAGTGTGAACCCCGTTGGAGGTATATCG-3'
<i>TaCBFIVa-A2</i>	A2_F1 A2_R1	Sens Anti-sens	5'-TCGTTACACCAACCAACCACTCTCAACG-3' 5'-GCACTAGCAGTACTCTCTAGATCTTTACATTGCTTA-3'
<i>TaCBFIVa-2.2</i>	2.2_F1 2.2_R1	Sens Anti-sens	5'-CCTCGTTACACCAACAACTGCTCTCAAC-3' 5'-AGCACTAGCAGTACTCTCTAGAACTTACATTGCTT-3'
<i>TaCBFIVa-25.1</i>	25.1_F1 25.1_R1	Sens Anti-sens	5'-CTCAGTCTCCGAAAGTCCGAAACACAC-3' 5'-AACACAGTACTGCTCTCTACAACTTACACATCC-3'
<i>TaCBFIVa-25.2</i>	25.2_F1 25.2_R1	Sens Anti-sens	5'-CTCCGGGCTCACTAGTCTGACACACAC-3' 5'-CAGTGTGCTGCTGCCCTCTACAACTTTAC-3'
<i>TaCBFIVb-A20</i>	A20_F1 A20_R1	Sens Anti-sens	5'-ATCCCAACTCCACCGCTCAAAACC-3' 5'-TCTTCCCATATTGGACGAACACAACC-3'
<i>TaCBFIVb-B20</i>	B20_F1 B20_R1	Sens Anti-sens	5'-ACTCAAAACCAACCTGCAACTCTCAAC-3' 5'-GTTTCTAAAACTGTACATGCCCTTAATTAGTCGGATA-3'
<i>TaCBFIVb-D20</i>	D20_F1	Sens	5'-ATCCCAACTCGACCGCTCAAAACC-3'



	D20 R1	Anti-sens	5'-GAGGGAGTAGTTCCAAAACGTGACATGGCTTA-3'
<i>TaCBFIVb-21.1</i>	21.1 F1	Sens	5'-CAATCCATCGACCCTCACCACAAAC-3'
	21.1 R1	Anti-sens	5'-GTGGGACAGAAAGTGCAGAGCCAACT-3'
<i>TaCBFIVb-21.2</i>	21.2 F1	Sens	5'-CACTCCATCGACCCTCAGCACAAAC-3'
	21.2 R1	Anti-sens	5'-TTTGGTGGACAGAAAGTACAGAACCAACT-3'
<i>TaCBFIVb-D21</i>	D21 F1	Sens	5'-GGAGCCCCAATCAATCGAACCTC-3'
	D21 R1	Anti-sens	5'-CCATATCTGGTGGGACAGAAAGTATAGAACCAAC-3'
<i>TaCBFIVc-14.1</i>	14.1 F1	Sens	5'-GCAGCAGCTAACTCCCTGCTTAATTACC-3'
	14.1 R1	Anti-sens	5'-TGGAACAGAAAGCAGAGAAACCGTCTAAAAAC-3'
<i>TaCBFIVc-B14</i>	B14 F1	Sens	5'-CTCAACGCGCAGCTAAACACGCTAAG-3'
	B14 R1	Anti-sens	5'-GGTGGAACAGAAAGAGAGAACTATCTATACAAACAT-3'
<i>TaCBFIVc-14.3</i>	14.3 F1	Sens	5'-CTCACCGCCCTCTCTCCAGCAT-3'
	14.3 R1	Anti-sens	5'-GGTGGAACAGAAAGAGAGAAACCATCTACAAAC-3'
<i>TaCBFIVd-4.1</i>	4.1 F1	Sens	5'-ACACCATCGCTCAACGTGCTCAAC-3'
	4.1 R1	Anti-sens	5'-TGGAAGAACACACACGCAACTATCGAC-3'
<i>TaCBFIVd-B4</i>	B4 F1	Sens	5'-AACACCATCGCTCAACCTGCTCAAC-3'
	B4 R1	Anti-sens	5'-CCAGATCGGGAAGAACACACAGAAC-3'
<i>TaCBFIVd-26.2</i>	26.2 F1	Sens	5'-CAAACTCTCTACACAGCTGCTGATCTTCC-3'
	26.2 R1	Anti-sens	5'-TGCTATACACAAACGAAAGAAACAACTGCCTACTCT-3'
<i>TaCBFIVd-9.1</i>	9.1 F1	Sens	5'-CAACACAGCCGTTGATTCTTCCACTACT-3'
	9.1 R1	Anti-sens	5'-AGGCGACTATCTACAACTTGACAGTATCGCTAA-3'
<i>TaCBFIVd-B9</i>	B9 F1	Sens	5'-GCCGCTGATTCCATCACTACTACCCTC-3'
	B9 R1	Anti-sens	5'-AAGAGCGGCACTATCTACAACTTGGCAGTA-3'
<i>TaCBFIVd-D9</i>	D9 F1	Sens	5'-CCGCTGATTCTTCCAGTACTCCTGCT-3'
	D9 R1	Anti-sens	5'-CTGCCTAGTATGTTTCTTCTCTAAATCTACAACTTGA-3'
<i>TaCBFIVd-A22</i>	A22 F1	Sens	5'-CCAGTTCTCAGCGGAACACAGCAAC-3'
	A22 R1	Anti-sens	5'-AACAAGATGCTATCTATGGTACAGCTTTACACTGC-3'
<i>TaCBFIVd-B22</i>	B22 F1	Sens	5'-CCTCCAGTTCTCAGCACACATCGAC-3'
	B22 R1	Anti-sens	5'-CGGACAGAGGGAGTACAACTTACACTGCTAGAT-3'

## RÉFÉRENCES

- Able JA, Langridge P, Milligan AS (2007) Capturing diversity in the cereals: many options but little promiscuity. *TRENDS in Plant Science*, 12: 71-79.
- Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M (1998) A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO Journal*, 17: 5484-5496.
- Arnal C, Girot B (2002) Drépanocytose chez adulte. *Encycl Med Chir (Elsevier)*. Traité d'hématologie, 13-006-D16 :15p.
- Artus NN, Uemura M, Steponkus PL, Gilmour SJ, Lin C, Thomashow MF (1996) Constitutive expression of the cold-regulated *Arabidopsis thaliana* *COR15a* gene affect both chloroplast and protoplast freezing tolerance. *Plant Biology*, 93: 13404-13409.
- Aung P, Shunji M, Kazuya K, Kyoko K, Kazuhiro Y, Keisuke M, Hirofumi N, Wataru Y (2008) No evidence of correlation between the single nucleotide polymorphism of DNMT3B promoter and gastric cancer risk in a Japanese population. *Oncology Reports*, 14: 1151-1154.
- Badawi M (2008) Fonctions génomiques de la famille des facteurs de transcription AP2 chez les céréales. Thèse de doctorat en biologie (Université du Québec à Montréal).

- Badawi M, Danyluk J, Boucho B, Houde M, Sarhan F (2007) The *CBF* gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. *Molecular Genetics and Genomics*, 277: 533-554.
- Badawi M, Venkat Reddy Y, Agharbaoui Z, Tominaga Y, Danyluk J, Sarhan F, Houde M (2008) Structure and functional analysis of wheat *ICE* (Inducer of CBF expression) genes. *Plant Cell Physiology*, 49: 1237-1249.
- Båga M, Chodaparambil V, Limin AE, Pecar M, Fowler DB, Chibbar R (2007) Identification of quantitative trait loci and associated candidate genes for low-temperature tolerance in cold-hardy winter wheat. *Genomics*, 7: 53-68.
- Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana cor15a* has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Molecular Biology*, 24: 701-713.
- Bhattaramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register JC, Tingey SV, Rafalski A (2002) Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Molecular Biology*, 48: 539-547.
- Bracale M, Coraggio I (2003) Chilling and freezing stresses in plants: Cellular responses and molecular strategies for adaptation. In: di Toppi LS, Abiotic stresses in plants. Kluwer Academic Pub: London pp: 23-52.
- Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates, tempo and mode of evolution. *Journal of Molecular Evolution*, 18: 225-239.
- Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP (1997) Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell Molecular and Life Science*, 53: 621-645



- Canella-D, Gilmour SJ, Kuhn LA, Thomashow MF (2010) DNA binding by the *Arabidopsis* CBF1 transcription factor requires the PKKP/RAGR<sub>x</sub>KKFxETRHP signature sequence. *Biochimica et Biophysica Acta*, 1799: 454-462.
- Cartwright B, Zarcinas BA, Mayfield AH (1984) Toxic concentrations of boron in a red-brown earth at Gladstone, South Australia. *Australian Journal of Soil Research*, 22: 261-272.
- Catala R, Salinas J (2008) Regulatory mechanisms involved in cold acclimation response. *Spanish Journal of Agricultural Research*, 6: 211-220.
- Chen HH, Li PH, Brenner ML (1983) Involvement of abscisic acid in potato cold acclimation. *Plant Physiology*, 71: 362-365.
- Chen Z, Cheng CHC, Zhang J, Cao L, Chen L, Zhou L, Jin Y, Ye H, Deng C, Dai Z, Zue Q, Hu P, Sun S, Shen Y (2008) Transcriptomic and genomic evolution under constant cold in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 12944-12949.
- Chinnusamy V, Zhu J, Zhu JK (2006) Gene regulation during cold acclimation in plants. *Physiologia Plantarum*, 126: 52-61.
- Chinnusamy V, Zhu J, Zhu JK (2007) Cold stress regulation of gene expression in plants. *TRENDS in Plant Science*, 12: 444-451.
- Choi DW, Rodriguez EM, Close TJ (2002) Barley *CBF3* gene identification, expression pattern, and map location. *Plant Physiology*, 129: 1781-1787.

- Choi EY, Kolesik P, McNeill A, Collins H, Zhang QS, Huynh BL, Graham RD, Stangoulis J (2007) The mechanism of boron tolerance for maintenance of root growth in barley (*Hordeum vulgare L.*). Plant, Cell and Environment, 30: 984-993.
- Cuesta-Marcos A, Szűcs P, Close TJ, Filichkin T, Muehlbauer GJ, Smith KP, Hayes PM (2010) Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in barley: implications for association mapping in germplasm arrays varying in size and structure. BMC Genomics, 11: 707-721.
- Curtis SE, Clegg MT (1984) Molecular evolution of chloroplast DNA sequences. Molecular Biology and Evolution, 1: 291-301.
- Danyluk J, Kane NA, Breton G, Limin AE, Fowler DB, Sarhan F (2003) TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. Plant Physiology, 132: 1849-1860.
- De Jong WW, Zweers A, Cohen LH (1978) Influence of single amino acid substitutions on electrophoretic mobility of sodium dodecyl sulfate-protein complexes. Biochemical and Biophysical Research Communications, 82: 532-539.
- Den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Human Mutation, 15: 7-12.
- Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell, 127: 1309-1321.
- Dubcovsky J, Chen C, Yan L (2005) Molecular characterization of the allelic variation at the *VRN-H2* vernalization locus in barley. Molecular Breeding, 15: 395-407.

- Dubreuil P, Dufour P, Krejci E, Causse M, De Vienne D, Gallais A, Charcosset A (1996) Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups. *Crop Science*, 36: 790-799.
- Dunn MA, White AJ, Vural S, Hughes MA (1998) Identification of promoter elements in a low-temperature-responsive gene (*blt4.9*) from barley (*Hordeum vulgare* L.). *Plant Molecular Biology*, 38: 551-564.
- Eyal Y, Meller Y, Lev Yadun S, Fluhr R (1993) A basic-type PR-1 promoter directs ethylene responsiveness, vascular and abscission zone-specific expression. *Plant Journal*, 4: 225-234.
- Fawcett AJ, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences of the United States of America*, 106: 5737-5742.
- Ferea TL, Botstein D, Brown PO, Rosenzweig RF (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 96: 9721-9726.
- Ferrari JA, Georgiou GP (1991) Quantitative genetic variation of esterase activity associated with a gene amplification in *Culex quinquefasciatus*. *Heredity*, 66: 265-272.
- Fowler DB, Breton G, Limin AE, Mahfoozi S, Sarhan F (2001) Photoperiod and temperature interactions regulate low-temperature-induced gene expression in barley. *Plant Physiology*, 127: 1676-1681.
- Fowler DB, Limin AE, Ritchie JT (1999) Low-temperature tolerance in cereals: model and genetic interpretation. *Crop Science*, 39: 626-633.



- Fowler S, Thomashow MF (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the *CBF* cold response pathway. *Plant Cell*, 14: 1675-1690.
- Francia E, Barabaschi D, Tondelli A, Laidò G, Rizza F, Stanca AM, Busconi M, Fogher C, Stockinger EJ, Pecchioni N (2007) Fine mapping of a *HvCBF* gene cluster at the frost resistance locus *Fr-H2* in barley. *Theoretical and Applied Genetics*, 115: 1083-1091.
- Francia E, Rizza F, Cattivelli L, Stanca AM, Galiba G, Tóth B, Hayes PM, Skinner JS, Pecchioni N (2004) Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) x 'Tremois' (spring) barley map. *Theoretical and Applied Genetics*, 108: 670-680.
- Fricano A, Rizza F, Faccioli P, Pagani D, Pavan P, Stella A, Rossini L, Piffanelli P, Cattivelli L (2009) Genetic variants of *HvCbf14* are statistically associated with frost tolerance in a European germplasm collection of *Hordeum vulgare*. *Theoretical and Applied Genetics*, 119: 1335-1348.
- Fu D, Szucs P, Yan L, Helguera M, Skinner J, Hayes P, Dubcovsky J (2005) Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics*, 273: 54-65.
- Gaboriaud C, Bissery V, Benchetrit T, Morron JP (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Letters*, 224: 149-155.
- Galiba G, Quarrie SA, Sutka J, Morgounov A, Snape JW (1995) RFLP mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. *Theoretical and Applied Genetics*, 90: 1174-1179.
- Galiba G, Vágújfalvi A, Li C, Soltész A, Dubcovsky J (2009) Regulatory genes involved in the determination of frost tolerance in temperate cereals. *Plant Science*, 176: 12-19.

- Gibson TA, Goldberg DS (2009) Questioning the ubiquity of neofunctionalization. *Plos Computational Biology*, 5: e1000252.
- Gill BS, Appels R, Botha Oberholster AM, Buell CR, Bennetzen JL, Chalhoub B, Chumley F, Dvolák J, Iwanaga M, Keller B, Li W, McCombie WR, Ogihara Y, Quetier F, Sasaki T (2004) A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium. *Genetics*, 168: 1087-1096.
- Gilmour SJ, Fowler SG, Thomashow MF (2004) *Arabidopsis* transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. *Plant Molecular Biology*, 54: 767-781.
- Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression. *The Plant Journal*, 16: 433-442.
- Goff SA, Ricke D, Lan TH, Presting T, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun W, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalima T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science*, 296: 92-100.
- Gojobori T, Li WH, Graur D (1982) Patterns of nucleotide substitution in pseudogenes and functional genes. *Journal of Molecular Evolution*, 18: 360-369.
- Gu X (2003) Evolution of duplicate genes versus genetic robustness against null mutations. *Trends in Genetics*, 19: 354-356.

- Guy CL (2003) Freezing tolerance of plants: current understanding and selected emerging concepts. *Canadian Journal of Botany*, 81: 1216-1223.
- Guy CL, Niemi KJ, Brambl R (1985) Altered gene expression during cold acclimation of spinach. *Proceedings of the National Academy of Sciences of the United States of America*, 82: 3673-3677.
- Hahn MW (2009) Distinguishing among evolutionary models for the maintenance of gene duplicates. *Journal of Heredity*, 100: 605-617.
- Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A (1999) Patterns of single nucleotide polymorphisms in candidate genes for blood pressure homeostasis. *Nature Genetics*, 22: 239-247.
- Hao D, Ohme Takagi M, Sarai A (1998) Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF Domain) in plant. *The Journal of Biological Chemistry*, 273: 26857-26861.
- Hao D, Yamasaki K, Sarai A, Ohme-Takagi M (2002) Determinants in the sequence specific binding of two plant transcription factors, *CBF1* and *NtERF2*, to the DRE and GCC motifs. *Biochemistry*, 41: 4202-4208.
- Hassan M, Oldach K, Baumann U, Langridge P, Sutton T (2010) Genes mapping to boron tolerance QTL in barley identified by suppression subtractive hybridization. *Plant, Cell and Environment*, 33: 188-198.
- Henderson IR, Shindo C, Dean C (2003) The needs for winter in the switch to flowering. *Annual Review of Genetics*, 37: 371-392.



- Houde M, Belcaid M, Ouellet F, Danyluk J, Monroy A, Dryanova A, Gulick P, Bergeron A, Laroche A, Links M, McCarthy L, Crosby WL, Sarhan F (2006) Wheat EST resources for functional genomics of abiotic stress. *BMC Genomics*, 7: 149.
- Huang D, Graham RD (1990) Resistance of wheat genotypes to boron toxicity is expressed at the cellular level. *Plant and Soil*, 126: 295-300.
- Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Research*, 9: 868-877.
- Iwaki K, Nishida J, Yanagisawa T, Yoshida H, Kato K (2002) Genetic analysis of *Vrn-B1* for vernalization requirement by using linked dCAPS marker in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 104: 571-576.
- Jaglo KR, Kleff S, Amundsen KL, Zhang X, Haake V, Zhang JZ, Deits T, Thomashow MF (2001) Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold response pathway are conserved in *Brassica napus* and other plant species. *Plant Physiology*, 127: 910-917.
- Jaglo Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF (1998) *Arabidopsis* *CBFI* overexpression induces *COR* genes and enhances freezing tolerance. *Science*, 280: 104-106.
- Jefferies S, Barr A, Gianquitto P, Warner C, Lewis J, Wheeler R (2001) Agronomic and breeding value of major genes in an introgression program. Proceedings of the 10th Australian Barley Technical Symposium, Canberra, ACT, Australia.
- Jefferies SP, Barr AR, Karakousis A, Kretschmer JM, Manning S, Chalmers KJ (1999) Mapping of chromosomal regions conferring boron toxicity tolerance in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, 98: 1293-1303.

- Jiang C, Iu B, Singh J (1996) Requirement of a CCGAC cis-acting element for cold induction of the *BN115* gene from winter *Brassica napus*. *Plant Molecular Biology*, 30: 679-684.
- Kanazin V, Talbert H, See D, DeCamp P, Nevo E, Blake T (2002) Discovery and assay of single-nucleotide polymorphisms in barley (*Hordeum vulgare*). *Plant Molecular Biology*, 48: 529-537.
- Karsai I, Szucs P, Meszaros K, Filichkina T, Hayes PM, Skinner JS, Lang L, Bedo Z (2005) The *Vrn-H2* locus is a major determinant of flowering time in a facultative x winter growth habit barley (*Hordeum vulgare* L.) mapping population. *Theoretical and Applied Genetics*, 110: 1458-1466.
- Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, 217: 624-626.
- Knox AK, Li C, Vágújfalvi A, Galiba G, Stockinger EJ, Dubcovsky J (2008) Identification of candidate *CBF* genes for the frost tolerance locus *Fr-Am2* in *Triticum monococcum*. *Plant Molecular Biology*, 67: 257-270.
- Knox AK, Dhillon T, Cheng H, Tondelli A, Pecchioni N, Stockinger EJ (2010) *CBF* gene copy number variation at *Frost Resistance-2* is associated with levels of freezing tolerance in temperate-climate cereals. *Theoretical and Applied Genetics*, 121: 21-35.
- Kochian LV (1995) Cellular mechanisms of aluminum toxicity and resistance in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 46: 237-260.
- Kochian LV, Hoekenga OA, Pineros MA (2004) How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorus efficiency. *Annual Review Plant Biology* 55: 459-493

- Kondrashov FA, Kondrashov AS (2006) Role of selection in fixation of gene duplications. *Journal of Theoretical Biology*, 239: 141-151.
- Limin AE, Danyluk J, Chauvin L-P, Fowler DB, Sarhan F (1997) Chromosome mapping of low-temperature induced *Wcs120* family genes and regulation of cold-tolerance expression in wheat. *Molecular and General Genetics*, 253: 720-727.
- Limin AE, Fowler DB (2006) Low-temperature tolerance and genetic potential in wheat (*Triticum aestivum* L.): response to photoperiod, vernalization, and plant development. *Planta*, 224: 360-366.
- Lin C, Tomashov MF (1992) DNA sequence analysis of a complementary DNA for cold-regulated *Arabidopsis* gene *cor15* and characterization of the COR 15 polypeptide. *Plant Physiology*, 99: 519-525.
- Liu J, Zhang Y, Lei X, Zhang Z (2008) Natural selection of protein structural and functional properties: a single nucleotide polymorphism perspective. *Genome Biology*, 9: R69.1-R69.17.
- Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature - responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell*, 10: 1391-1406.
- Lundqvist U, Franckowiack JD (1997) BGS 6, six-rowed spike 1, *vrsl*. *Barley Genetic Newsletter* 26: 49-50.
- Mahfoozi S, Limin AE, Hayes PM, Hucl P, Fowler DB (2000) Influence of photoperiod response in the expression of cold hardness in wheat and barley. *Canadian Journal of Plant Science*, 80: 721-724.



- Mao Y, Pavangadkar KA, Thomashow MF, Triezenberg SJ (2006) Physical and functional interactions of *Arabidopsis* ADA2 transcriptional coactivator proteins with the acetyltransferase GCN5 and with the cold-induced transcription factor CBF1. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1759: 69-79.
- Matsumoto H (2000) Cell biology of aluminum toxicity and tolerance in higher plants. *International Review of Cytology*, 200: 1-46.
- Matsumoto H (2000) Cell biology of aluminum toxicity and tolerance in higher plants. *International Review of Cytology*, 200: 1-46.
- McKhann H, Gery C, Bérard A, Lévêque S, Zuther H, Hinch DK, De Mita S, Brunel D, Téoulé E (2008) Natural variation in *CBF* gene sequence, gene expression and freezing tolerance in the Versailles core collection of *Arabidopsis thaliana*. *BMC Plant Biology*, 8: 105.
- Miles C, Wayne M (2008) Quantitative trait locus (QTL) analysis. *Nature Education*, 1 (1).
- Miller AK, Galiba G, Dubcovsky J (2006) A cluster of 11 CBF transcription factors is located at the frost tolerance locus *Fr-A<sup>m</sup>2* in *Triticum monococcum*. *Molecular Genetics and Genomics*, 275: 193-203.
- Moore RC, Grant SR, Purugganan MD (2005) Molecular population genetics of redundant floral-regulatory genes in *Arabidopsis thaliana*. *Molecular Biology and Evolution*, 22: 91-103.
- Murai K, Miyamae M, Kato H, Takumi S, Ogiwara Y (2003) *WAP1*, a wheat *APETALA1* homolog, plays a central role in the phase transition from vegetative to reproductive growth. *Plant Cell Physiology*, 44: 1255-1265.

- Nable RO (1988) Effects of B toxicity amongst several barley wheat cultivars: a preliminary examination of the resistance mechanism. *Plant Soil*, 112: 45-52.
- Novillo F, Alonso JM, Ecker JR, Salinas J (2004) CBF2/DREB1C is a negative regulator of *CBF1/DREB1B* and *CBF3/DREB1A* expression and plays a central role in stress tolerance in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 3985-3990.
- Novillo F, Medina J, Salinas J (2007) *Arabidopsis* CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. *Proceedings of the National Academy of Sciences of the United States of America*, 104: 21002-21007.
- Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM (2004) Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. *Molecular Biology and Evolution*, 21: 1308-1317.
- Oerti JJ, Roth JA (1969) Boron supply of sugar beet, cotton and soybean. *Agronomy Journal*, 61: 191-195.
- Ohme-Takagi M, Shinshi H (1990) Structure and expression of a tobacco  $\beta$ -1,3-glucanase gene. *Plant Molecular Biology*, 15: 941-946.
- Ohno S (1970) *Evolution by gene duplication*. Springer, New York.
- Palaisa K, Morgante M, Tingey S, Rafalski A (2004) Long range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 9885-9890.

- Phillips JR, Dunn MA, Hughes MA (1997) mRNA stability and localisation of the low-temperature-responsive barley gene family *blt14*. *Plant Molecular Biology*, 33: 1013-1023.
- Rebourg C, Dubreuil P, Charcosset A (1999) Genetic diversity among maize populations: bulk RFLP analysis of 65 accessions. *Maydica*, 44: 237-249.
- Reid R (2007) Identification of boron transporter genes likely to be responsible for tolerance to boron toxicity in wheat and barley. *Plant Cell Physiology*, 48: 1673-1678.
- Riechmann JL, Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. *Biological Chemistry*, 379: 633-646.
- Rifkin SA, Kim J, White KP (2003) Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nature Genetics*, 33: 138-144.
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Bioinformatics Methods and Protocols*, 132: 365-386.
- Sakuma Y, Liua Q, Dubouzeta JG, Abea H, Shinozaki K, Yamaguchi-Shinozaki K (2002) DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration and cold-induced gene expression. *Biochemical and Biophysical Research Communications*, 290: 998-1009.
- Sandve SR, Fjellheim S (2010) Did gene family expansions during the Eocene-Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates? *Molecular Ecology*, 19: 2075-2088.



- Sarhan F, Ouelett F, Vazquez-Tello A (1997) The wheat *wcs120* gene family. A useful model to understand the molecular genetics of freezing tolerance in cereals. *Physiologia Plantarum*, 101: 439-445.
- Sasaki T, Ryan PR, Delhaize E (2006) Sequence upstream of the wheat (*Triticum aestivum* L.) *ALMT1* gene and its relationship to aluminum resistance. *Plant and Cell Physiology*, 47: 1343-1354.
- Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell*, 13: 61-72.
- Sharma P, Sharma N, Deswal R (2005) The molecular biology of the low-temperature response in plants. *BioEssays*, 27: 1048-1059.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, 29: 308-311.
- Shinozaki K, Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinion in Plant Biology*, 3: 217-223.
- Shinozaki-Yamaguchi K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends Plant Science*, 10: 88-94.
- Skinner JS, Szucs P, von Zitzewitz J, Marquez-Cedillo L, Filichkin T, Stockinger EJ, Thomashow MF, Chen TH, Hayes PM (2006) Mapping of barley homologs to genes that regulate low temperature tolerance in *Arabidopsis*. *Theoretical and Applied Genetics*, 112: 832-842.

- Skinner JS, Zitzewitz JV, Szucs P, Marquez-Cedillo L, Filichkin T, Amundsen K, EJ Stockinger, Thomashow MF, Chen THH, Hayes PMM (2005) Structural, functional, and phylogenetic characterization of a large *CBF* gene family in barley. *Plant Molecular Biology*, 59: 533-551.
- Smith MJ, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetical Research*, 23: 23-25.
- Snape JW, Semikhodskii A, Sarma R, Korzun V, Fish L, Quarrie SA, Gill BS, Sasaki T, Galiba G, Sutka J (1998) Mapping vernalization loci in wheat and comparative mapping with other cereals. *Proceedings of the 9th International Wheat Genetics Symposium*, 3: 156-158.
- Snape JW, Sarma R, Quarrie SA, Fish L, Galiba G, Sutka J (2001) Mapping genes for flowering time and frost tolerance in cereals using precise genetic stocks. *Euphytica*, 120: 309-315.
- Somers DJ, Kirkpatrick R, Moniwa M, Walsh A (2003) Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. *Genome*, 49: 431-437.
- Steponkus PL, Uemura M, Joseph RA, Gilmour SG, Thomashow MF (1998) Mode of action of the *COR15a* gene on the freezing tolerance of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, 95: 14570-14575.
- Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana* *CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Plant Biology*, 94: 1035-1040.

- Stockinger EJ, Mao Y, Regier MK, Triezenberg ST, Thomashow MF (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. *Nucleic Acids Research*, 29: 1524-1533.
- Stockinger EJ, Skinner JS, Gardner KG, Francia E, Pecchioni N (2007) Expression levels of barley *CBF* genes at the *Frost resistance-H2* locus are dependent upon alleles at *Fr-H1* and *Fr-H2*. *The Plant Journal*, 51: 308-321.
- Straub PF, Shen Q, Ho THD (1994) Structure and promoter analysis of an ABA and stress-regulated barley gene, *HVA1*. *Plant Molecular Biology*, 26: 617-630.
- Sung DY, Kaplan F, Lee KJ, Guy CL (2003) Acquired tolerance to temperature extremes. *TRENDS in Plant Science*, 8: 179-187.
- Sutton T, Baumann U, Hayes J, Collins NC, Shi B-J, Schnurbusch T (2007) Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science*, 318: 1446-1449.
- Takumi S, Koike A, Nakata M, Kume S, Ohno R, Nakamura C (2003) Cold-specific and light-stimulated expression of a wheat (*Triticum aestivum* L.) *Cor* gene *Wcor15* encoding a chloroplast-targeted protein. *Journal of Experimental Botany*, 54: 2265-2274.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28: 2731-2739.
- Tenaillon MI, Swakins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays ssp. mays* L.). *Proceedings of the National Academy of Sciences of the United States of America*, 98: 9161-9166.



- Thomashow MF (1998) Role of cold-responsive genes in plant freezing tolerance. *Plant Physiology*, 118: 1-7.
- Thomashow MF (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology*, 50: 571-599.
- Thomashow MF (2001) So what's new in the field of plant cold acclimation? Lots! *Plant Physiology*, 25: 89-93.
- Toth B, Galiba G, Fehér E, Sutka J, Snape W (2003) Mapping genes affecting flowering time and frost resistance on chromosome 5B of wheat. *Theoretical and Applied Genetics*, 107: 509-514.
- Trevaskis B, Hemming MN, Peacock WJ, Dennis ES (2006) *HvVRN2* responds to daylength, whereas *HvVRN1* is regulated by vernalization and developmental status. *Plant Physiology*, 140: 1397-1405.
- Vágújfalvi A, Aprile A, Miller A, Dubcovsky J, Delugu G, Galiba G, Cattivelli L (2005) The expression of several *CBF* genes at the *Fr-A2* locus is linked to frost resistance in wheat. *Molecular Genetics and Genomics*, 274: 506-514.
- Vágújfalvi A, Crosatti C, Galiba G, Dubcovsky J, Cattivelli L (2000) Two loci on wheat chromosome 5A regulate the differential cold-dependent expression of the *cor14b* gene in frost tolerant and sensitive genotypes. *Molecular Genetics and Genomics*, 263: 194-200.
- Vágújfalvi A, Galiba G, Cattivelli L, Dubcovsky J (2003) The cold-regulated transcriptional activator Cbf3 is linked to the frost tolerance locus *Fr-A2* on wheat chromosome 5A. *Molecular Genetics and Genomics*, 269: 60-67.

- Wakeley J (1994) Substitution-rate variation among sites and the estimation of transition bias. *Molecular Biology and Evolution*, 11: 436-442.
- Wakeley J (1996) The excess of transitions among nucleotide substitutions: New methods of estimating transition bias underscore its significance. *Trends in Ecology and Evolution*, 11: 158-163.
- Wallis JG, Wang H, Guerra DJ (1997) Expression of a synthetic antifreeze protein in potato reduces electrolyte release at freezing temperatures. *Plant Molecular Biology*, 35: 323-330.
- Wang W, Spurdle AB, Kolachana P, Bove B, Modan B, Ebberts SM, Suthers G, Tucker MA, Kaufman DJ, Doody MM, Tarone RE, Daly M, Levavi H, Pierce H, Chetrit H, Godwin AK, Struwing JP (2001) Single nucleotide polymorphism in the 5' untranslated region of RAD51 and risk of cancer among BRCA1/2 mutation carriers. *Cancer Epidemiology, Biomarkers and Prevention*, 10: 955-960.
- Wang Z, Triezenberg SJ, Thomashow MF, Stockinger EJ (2005) Multiple hydrophobic motifs in *Arabidopsis* CBF1 COOH-terminus provide functional redundancy in trans-activation. *Plant Molecular Biology*, 58: 543-559.
- White AJ, Dunn AM, Brown K, Hughes MA (1994) Comparative analysis of genomic sequence and expression of a lipid transfer protein gene family in winter barley. *Journal of Experimental Botany*, 45: 1885-1892.
- Wilens RW, Fu P, Robertson AJ (1996) A comparison of the cold hardiness potential of spring cereals and vernalized and non-vernalized winter cereals. In: Li PH, Chen THH (eds) Fifth international plant cold hardiness seminar. Plenum Press, Corvallis, pp 191-201.
- Wise MJ, Tunnacliffe A (2004) POPP the question: what do LEA proteins do? *TRENDS in Plant Science*, 9: 13-17.

- Xiao FH, Xue GP (2001) Analysis of promoter activity of late embryogenesis abundant protein genes in barley seedlings under conditions of water deficit. *Plant Cell Reports*, 20: 667-673.
- Xiong L, Schumaker KS, Zhu JK (2002) Cell signaling during cold, drought, and salt stress. *The Plant Cell*, 14: S165-S183.
- Xue G (2002) An AP2 domain transcription factor HvCBF1 activates expression of cold-responsive genes in barley through interaction with a (G/a)(C/t) CGAC motif. *Biochimica et Biophysica Acta*, 1577: 63-72.
- Xue G (2003) The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature. *The Plant Journal*, 33: 373-383.
- Yamada T, Kuroda K, Jitsuyama Y, Takezawa D, Arakawa K, Fujikawa S (2002) Roles of the plasma membrane and the cell wall in the responses of plant cells to freezing. *Planta*, 215: 770-778.
- Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel cis acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high salt stress. *Plant Cell*, 6: 251-264.
- Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JF, Echenique V, Dubcovsky J (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science*, 303: 1640-1644.
- Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences of the United States of America*, 100: 6263-6268.



- Zarka DG, Vogel JT, Cook D, Thomashow MF (2003) Cold induction of *Arabidopsis* *CBF* genes involves multiple ICE (inducer of *CBF* expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. *Plant Physiology*, 133: 910-918.
- Zou C, Lehti-Shiu MD, Thomashow M, Shiu SH (2009) Evolution of stress-regulated gene expression in duplicate genes of *Arabidopsis thaliana*. *PLoS Genetics*, 5: e1000581.